Examining the Role of Urban Spatial Structure, Housing Submarkets, and Economic Resiliency in U.S. Residential Foreclosures, 2000-2009

by

Indro Ray

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2012 by the
Graduate Supervisory Committee:

Subhrajit Guhathakurta, Chair
Sergio Rey
Rhonda Phillips

ARIZONA STATE UNIVERSITY

May 2012

ABSTRACT

After a relative period of growth (2000-06), the U.S. economy experienced a sharp decline (2007-09) from which it is yet to recover. One of the primary factors that contributed to this decline was the sub-prime mortgage crisis, which triggered a significant increase in residential foreclosures and a slump in housing values nationwide. Most studies examining this crisis have explained the high rate of foreclosures by associating it with socio-economic characteristics of the people affected and their financial decisions with respect to home mortgages. Though these studies were successful in identifying the section of the population facing foreclosures, they were mostly silent about region-wide factors that contributed to the crisis. This resulted in the absence of studies that could identify indicators of resiliency and robustness in urban areas that are affected by economic perturbations but had different outcomes. This study addresses this shortcoming by incorporating three concepts. First, it situates the foreclosure crisis in the broader regional economy by considering the concept of regional economic resiliency. Second, it includes the concept of housing submarkets, capturing the role of housing market dynamics in contributing to market performance. Third, the notion of urban growth pattern is included in an urban sprawl index to examine whether factors related to sprawl could partly explain the variation in foreclosures. These, along with other important socio-economic and housing characteristics, are used in this study to better understand the variation in impacts of the current foreclosure crisis. This study is carried out for all urban counties in the U.S. between 2000 and 2009. The associations between foreclosure rates and different variables are established using spatial regression models. Based on these models, this dissertation argues that counties with higher degree of employment diversity, encouragement for small business enterprises, and with less dependence on housing related industries, experienced fewer foreclosures. In addition, this thesis concludes that the spatial location of foreclosed properties is a function of

location of origination of sub-prime mortgages and not the spatial location of the properties per se. Also importantly, the study found that the counties with high number of dissimilar housing submarkets experienced more foreclosures.

DEDICATION

I dedicate my dissertation to my late father, Kamalesh Ray. I wish he was here to see me reach this important milestone in my life.

ACKNOWLEDGMENTS

I would like to express my deep gratitude towards my mentor and dissertation committee chair Dr. Subhrajit Guhathakurta for his continuous support and encouragement throughout my doctoral study. He played a significant role in guiding me in articulating various aspects of my dissertation. He pushed me intellectually to work hard, to take control of my research and to produce high quality work. For these reasons and more, I will be ever grateful to him. I would also like to thank Dr. Sergio Rey for his insights and for introducing me to new research methods. He played a significant role, especially regarding the housing submarket part of this research. Additionally, I would like to thank Dr. Rhonda Philips for her insights. She encouraged me to introduce the concept of resiliency to my study, making regional economic resiliency a major part of this dissertation in its final form. I once again want to express my appreciation to all my committee members for their insights, support, and encouragement.

In addition to my committee members, I would like to thank Dr. Nabil Kamel, Mr. David Folch, and Mr. Tom Snyder for their help at different stages of my dissertation. I am thankful to my friends for their solidarity, and especially to Samantha for her support and encouragement throughout the course of my research. I would also like to convey my appreciation towards all the professors and administrative staff members involved in the Ph.D. program at the Herberger Institute for Design and the Arts and School of Geographical Sciences and Urban Planning for helping me throughout the course of the program. Finally, I express my gratitude to my mother, Ranu Ray for her unconditional love, without which I wouldn't be able to pursue and finish my doctoral study.

TABLE OF CONTENTS

Page

LIST OF TABLES vii

LIST OF FIGURES viii

CHAPTER

1 INTRODUCTION 1

Prologue to Foreclosure Crisis 1

Introduction to the Study 2

Organization of the Study 5

2 LITERATURE REVIEW 7

Introduction 7

Housing Market Crisis and Residential Foreclosures 7

Gaps in Current Literature 16

Economic Resiliency and its Measures 18

The Concept of Housing Submarket 23

Urban Sprawl and its Measure 28

3 METHODOLOGY 31

Conceptual Framework 31

Data Collection 32

Research Procedure 48

Expected Results 51

4 RESULTS AND ANALYSIS 53

Results 53

Analysis 64

CHAPTER Page

5 CONCLUSION 72

Conclusions 72

Contribution of the Study 74

Planning Implications 76

Limitations of the Study 78

Future Work and Directions 81

REFERENCES 83

LIST OF TABLES

Table		Page
1.	Variables used to construct urban sprawl index	39
2.	Descriptive statistics and data sources for the variables	47
3.	OLS Regression Diagnostics	55
4.	Diagnostics for Spatial Dependence in OLS Models	56
5.	Measures of fit for Ordinary Least Square and Spatial Error Model	57
6.	OLS and SEM results (2,000 HU submarket)	59
7.	OLS and SEM results (2,500 HU submarket)	60
8.	OLS and SEM results (3,000 HU submarket)	61
9.	OLS and SEM results (3,500 HU submarket)	62
10.	Diagnostics for Heteroskedasticity	63
11.	Diagnostics for Spatial Dependence	64
12.	Correlation between small firms growth rate for boom and bust periods	66
13.	Correlation between sprawl index and housing submarkets	81

LIST OF FIGURES

Figure		Page
1.	National seasonally adjusted housing purchase index (1991-2009).....	1
2.	Foreclosure rates in U.S counties (2009)	34
3.	Entropy Index in U.S counties (2006).....	36
4.	Sprawl index for urban counties (2000).....	39
5.	Housing submarkets in Maricopa County, Arizona.....	45
6.	Housing submarkets in Cook County, Illinois	46
7.	Research methodology flow diagram.....	52

CHAPTER 1

INTRODUCTION

Prologue to Foreclosure Crisis

During the last decade, the U.S. housing market went through a cycle of rapid growth and sharp decline in home values (see Figure 1). Studies have shown that the U.S. housing market has grown exponentially since 2000 and it reached its peak in mid-2006 (Case and Quigley, 2009; Mulligan, 2009). But since 2007, the housing prices have dropped, and together with a weak economy have resulted in an economic and housing crisis. This crisis has been marked by high rates of mortgage delinquency and foreclosure.

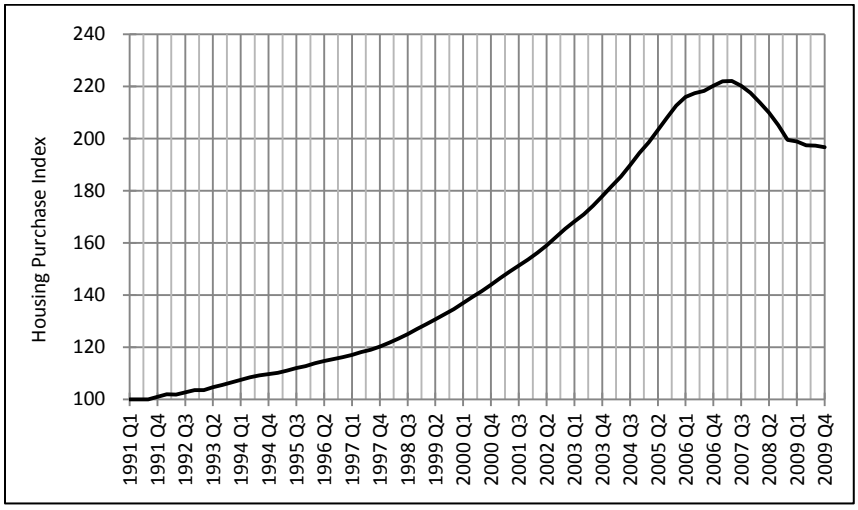

Figure 1. National seasonally adjusted housing purchase index (1991-2009) (1991 Q1=100); Source: Department of Housing and Urban Development, 2011.

A majority of the studies have concluded that the large number of high-risk subprime loans disbursed between 2000 and 2006 is among the major factors responsible for this crisis (Calem, Gillen, and Wachter, 2004; Gerardi and Willen, 2008; Goldstein, 2008; Mayer and Pence, 2008). Since 2000, easy access to high-risk loans contributed to rising housing prices and created a self-perpetuating cycle (Herbert, 2010; Hendershott, Hendershott, and Shilling, 2010). The incentives favored the mortgage

brokers who profited from a high volume of loans with exotic terms and loose underwriting controls and standards. Given that these mortgages were quickly bundled and sold to various banks and investment houses, the risks of default were transferred to different entities and were not a concern for the brokers. When the price increases finally slowed by mid-2006 together with increasing interest rates, consumers were suddenly faced with high interest payments on their mortgages. Homeowners could no longer refinance their mortgages when housing prices kept dropping and lending markets dried out. People owed more than their home's worth and the weakening economy resulted in loss of jobs. These circumstances resulted in high rates of mortgage default and subsequent foreclosures.

Introduction to the Study

After the mortgage lending market dried up in mid-2006, the housing values in the U.S. started to drop sharply. Combined with a weak housing market and a deteriorating economy, the U.S. faced high rates of unemployment and residential foreclosures. Between 2006 and 2009, the average annual number of foreclosure proceedings jumped from 275,000 to 1.7 million, a six-fold increase. A similar trend is projected for the next 2 years adding another 4 to 5 million foreclosures to already 6 million foreclosed properties (Davis, 2010). This housing crisis has attracted a considerable amount of attention from various academic arenas including planning. There are many studies conducted in the field of planning to understand the foreclosure problem and its effects on people and neighborhoods. Studies by Grover, Laura, and Richard (2007), Immergluck (2008), Mueller (2006), Ong and Pfeiffer (2008), and Pedersen and Delgadillo (2007) looked at the racial composition, educational attainment, household income levels, and unemployment rates among the people most affected by the housing crisis. Additionally, Bostic and Kwan (2008), Calem, Gillen, and Wachter (2003), and Immergluck (2009) looked into the housing and mortgage attributes. The

methods and the methodologies employed by the above studies are similar. They identified various social, economic, demographic, and housing variables and then, using statistical methods, commented on their possible association with high rates of foreclosure. This approach is useful for a basic diagnosis of the foreclosure crisis and to comment on certain attributes associated with the people facing the crisis. But this is not sufficient to have a deeper understanding of the problem in a broader context. The previous studies fall short for the following reasons.

First, the existing studies about the recent housing crisis have tried to understand the vulnerability of people and places towards the crisis using basic socio-economic indicators. They fail to provide an in-depth analysis of the crisis while situating the problem in a larger urban context. One of the common factors considered in the existing studies is the role of economic indicators in the foreclosure crisis. Authors have mainly used unemployment rate, change in unemployment rate over time, and percent self employed to establish association between the economy and the foreclosure crisis (Edmiston, 2009; Grover, Smith, and Todd, 2007; Immergluck and Smith, 2005; Immergluck, 2008). Though these indicators are good to draw and understand a generalized picture of the recent crisis, they do not provide sufficient information to draw specific conclusions. This study addresses this gap by considering the concept of regional economic resiliency. This concept is measured through different indicators such that it not only provides a broad view of an area's economy and its relationship to the crisis, but also point towards the elements in the economy that make a place more or less vulnerable to such housing crises. The variables used to measure the concept of economic resiliency at the regional level include workforce diversity, industrial competitiveness, growth of small and large business enterprises, and unemployment rate.

Second, the existing studies usually consider some property and neighborhood level housing characteristics to understand the association between housing market and housing foreclosures (Bostic and Kwan, 2008; Calem, Gillen, and Wachter, 2003; and Immergluck, 2009). These attributes include housing type, occupancy status, tenure status, age of housing, and median housing value. These attributes have been used and studied separately in the context of the recent crisis and this approach is useful in identifying and commenting on certain housing characteristics that make some areas more prone towards the crisis. But by using this approach, the existing studies fail to have a comprehensive look at the relation between housing market and the foreclosure crisis. This research addresses this shortcoming by considering the concept of housing submarket. This concept is considered because it has been argued as a good analytical framework to capture the housing market dynamics of an area (Jones and Watkins, 2009). Past housing studies (Goodman, 1978; Goodman, 1981; Maclennan and Tu, 1996; Bourassa, Hamelink, Hoesli, and MacGregor, 1999) indicate that local housing markets are a function of interactions among various structural and neighborhood characteristics. Some of these characteristics are considered in this study including median housing value, housing type, housing size, distance from work centers, and occupancy and tenure status. This study uses a new spatial clustering method to consider these factors altogether along with spatial contiguity of neighborhoods to form housing submarkets in each urban county in the U.S.

Third, the concept of urban sprawl has been studied extensively in the past because of its role in recent urban development and its impacts on various social, environment, and economic issues. In the recent studies by Immergluck (2010), Kaplan and Sommers (2009), Lehnert and Grover (2008), Leinberger (2008), Mayer and Pence (2008), Ong and Pfeiffer (2008), the authors have tried to incorporate this concept while looking into the factors contributing towards the recent foreclosure crisis in the U.S. In these studies, urban sprawl has been measured through various methods including travel

time to work, distance from central business district (CBD), and using dummy variables for certain neighborhoods based on their distance from the central city. But these studies did not consider a comprehensive measure of sprawl as it has been noted in the past that urban sprawl is more than distance from CBD (Ewing, Pendall, and Chen, 2002; Ewing, Schmid, Killingsworth, Zlot, and Raudenbush, 2003; Song and Knaap, 2004). Sprawl includes factors like population density, road density, growth pattern, worker's travel pattern, and distance from CBD among many others. This gap in the literature is filled in this research as it considers multiple measures and calculates a composite index of urban sprawl while looking into its possible association with the housing foreclosures.

Organization of the Study

This section discusses the organization of this document. The first chapter provides some background information on the current housing crisis along with the key issues that are addressed in this research. This is followed by the chapter reviewing the past literature pertaining to foreclosure crisis at the national, regional, and local levels. This chapter also identifies the significant socio-economic, housing, and spatial variables that have been associated with higher or lower rates of foreclosure. This analysis is followed by identifying the gaps in the current body of knowledge related to the foreclosure crisis and subsequently delves into the relevant literature. The third chapter of this document discusses the methodology adopted for this research and the methods used to analyze the data. This chapter starts with the various hypotheses set for this study followed by the scope of the research and descriptions of various data sets and data sources. It also identifies the dependent variable and multiple explanatory variables along with the processes of formulating them. This is followed by a discussion about the various statistical and spatial analysis methods used to analyze the data so that correct assessments can be made about the associations between foreclosure rate and the explanatory variables. The fourth chapter reports on the results and findings from

analysis of the data. It starts with the comparison of results between ordinary least square (OLS) and spatial regression models and subsequently analyzes the results of the spatial regression models that are deemed to be superior to OLS in this context. This analysis focuses on the connection between economic resiliency, urban sprawl, and nature of housing submarkets to foreclosure rate. The final chapter summarizes the findings and the conclusions drawn from the analysis. It is followed by sections about the contribution of this study, its limitations, and discussion of some ideas about future research directions.

CHAPTER 2

LITERATURE REVIEW

Introduction

This chapter discusses some of the important studies examining the recent housing market crisis with specific focus on residential foreclosures. For this thesis, the review of literature has been divided into two parts. The first part looks at the fundamental causes behind the recent crisis at the national level, while the second part outlines the factors contributing to high rates of foreclosures at the regional and neighborhood levels. Under each category, significant variables used in the past studies are discussed, along with the important findings. The purpose behind such analysis of the current literature is two-fold. First, it will provide a survey of methodology and data applied, together with the findings related to trends in housing foreclosure. Second, it will help to identify critical issues that may not show up in most studies. The latter purpose will be used as a guide to look into other relevant literature.

Housing Market Crisis and Residential Foreclosures

The current foreclosure crisis has attracted a considerable amount of attention from many academic fields. One of the main reasons for this attention is the magnitude of the crisis. The average foreclosure proceedings increased from 275,000 per year between the early 1980s and 2006 to 1.7 million between 2006 and 2009 (Davis, 2010). Related literature in the field of planning has extensively studied the causes and consequences of this crisis. The studies that examine the causes behind the crisis at the national level mainly deal with policies and regulations, whereas at the regional level studies look at the socio-economic, housing, and mortgage characteristics of households and neighborhoods, and their spatial attributes to explain foreclosures. The following sections cover these aspects with common and confounding findings across various

studies. Although there is extensive research available in the area of foreclosure crisis, several gaps in the literature still remain, which are also discussed in this chapter.

Foreclosure crisis at the national level.

The current crisis was mainly caused by high rates of mortgage delinquency that resulted in unprecedented number of foreclosures. Both of these events were preceded by a large number of high-risk loans disbursed between 2000 and 2006. The mortgage brokers were pushing such loans with exotic terms and loose underwriting controls, given that they benefited from the rewards but assumed none of the risks. Many households either did not understand the terms and conditions attached to them or they expected the market to continue to climb indefinitely, thereby nullifying the effects of future rate hikes.

The initial event that started the current economic and housing crises was the drop in housing prices, not a weak economy. In the early years of the crisis (mid 2006-2007), people avoided foreclosures by refinancing their loans or selling their property for profit. But the rise in home prices and availability of the high interest rate mortgages were linked, so the continuous drop in housing prices since 2007 was a sign that lenders could no longer keep lending money.

The increased availability of high-risk loans started in the 1980s and 1990s when legislative changes removed the interest rate ceiling on mortgages. This development was complemented by the use of statistical models and credit scores to estimate borrower risk and a shift in mortgage finance sources from federally regulated institutions to mortgage banking institutions that were subject to less regulatory oversight. Mayer, Pence, and Sherlund (2009) also argue that the regulatory structure to protect borrowers, such as the disclosure rules and the Home Ownership and Equity Protection Act (HOEPA), was inadequate to cover all the loans given away by the lending institutions. In addition to these loopholes, the rating agencies also played a significant

role in the current crisis. These agencies were optimistic about the risk associated with subprime mortgages and thus gave AA or AAA (the highest ratings for securities, bonds, and loans to invest based on predicted risk and investment returns) investment ratings to these loans (Herbert, 2010).

The above mentioned changes in policy and regulatory shortcomings set the stage for the current crisis. With inadequate underwriting controls, the number of subprime and Alt-A (loans requiring minimum or no asset or income documentation) loans saw a rapid increase around 2003. One of the major reasons behind the hike was the demand for high-yield, investment-grade securities from both foreign and domestic investors. This was complemented by the high profits associated with every stage of subprime loans, and most importantly, the rapid increase in housing prices through 2006 (Hendershott, Hendershott, and Shilling, 2010). Since 2000, the easy accessibility to high-risk loans contributed to the rising housing prices and created a self-perpetuating cycle. And when the prices finally slowed down by mid-2006, the true nature of the loans was exposed.

Most authors concur with the underlying causes behind the current foreclosure crisis noted above. A majority of the studies have blamed financial deregulation, subprime lending institutions, and credit rating agencies for the crisis. But there exists some disagreement about the role of government institutions and policies. Some authors have also been critical of the roles played by the Community Reinvestment Act (CRA) and Government Sponsored Enterprises (GSEs) in the current crisis.

To encourage home ownership and better quality of housing among low and middle-income families, the Community Reinvestment Act (CRA) was passed in 1977. Under this Act, local branches of banks are encouraged to meet the credit needs of low and middle-income families. Critics of CRA have argued that because of this Act, more high-risk loans were generated as the banks wanted to meet the CRA requirements. On the other hand, the supporters of the CRA suggest that a very small percentage of loans

given out under this program are high-risk loans. Furthermore, the CRA has a lending history of over 30 years whereas the foreclosure crisis is a recent event. Herbert (2010) also discusses that there is evidence that the loans given out under this Act performed better than the subprime mortgages during the crisis.

The critics who questioned CRA's role also suggest that government sponsored enterprises (GSEs) such as Fannie Mae and Freddie Mac, played a major role in promoting high-risk loans among low and middle-income households. There is a certain amount of truth to this claim as GSEs were purchasing a significant share of subprime mortgage backed securities to meet their low and middle-income housing needs. But there is evidence that suggests that these actions by GSEs were mainly driven by profit and market share, rather than fulfilling the housing needs. The data indicate that in 2004, half of their purchases were subprime and in the subsequent years, this percentage fell, whereas subprime mortgage rates boomed in these years (Herbert, 2010). Thus, it can be concluded that even though GSEs contributed to the rise in subprime mortgages, their action was motivated by the demand for securities from various investors.

Foreclosure crisis at the regional and local level.

The recent studies conducted to explain the causes behind the current foreclosure crisis can be broadly divided into two groups. The first group looked at the national level government policies and financial structure, while the second used various socio-economic, housing, and mortgage attributes to explain foreclosures at the local and regional levels. This section will cover the common attributes that different authors have found to have contributed towards high rates of foreclosure. It also discusses the results that are contradictory to the popular notion.

Socio-economic characteristics.

The studies related to foreclosure crisis in the field of planning have extensively used demographic and economic characteristics to explain variance in foreclosure rates across regions. One of the most common attributes selected is the minority status of the population (Pedersen and Delgadillo, 2007; Perkins, 2009; Mueller ,2006; and Grover, Laura, and Richard, 2007). These studies found that minority populations are disproportionately affected by the crisis. Some of the studies (Goldstein, 2008; Furman Center for Real Estate and Urban Policy, 2009) that used other variations of the same indicator such as proportion of minority homeownership and non-white mortgage borrowers also found similar results. Social characteristics like linguistic isolation and the number of households with a single female head are found to be associated with high foreclosure rates (Pedersen and Delgadillo, 2007; Perkins, 2009; Mueller, 2006). In terms of the educational attainment level, most researchers have studied the relationship between the foreclosure rate of an area and the proportion of the population with a college education (Grover, Laura, and Richard, 2007; Perkins, 2009; Ong and Pfeiffer, 2008). In all these cases, foreclosure rate had a negative association with college education. The variables related to the economic characteristics of the people included income level, financial health, and employment. Household income was found to be negatively associated with foreclosure (Pedersen and Delgadillo, 2007; Grover, Laura, and Richard, 2007; Mueller, 2006). In terms of employment characteristics, the unemployment rate is positively linked to the foreclosure rate (Immergluck, 2008; Grover, Laura, and Richard, 2007), while self-employment has a negative correlation with foreclosure (Pedersen and Delgadillo, 2007).

The above section identified the most common variables used by different authors to explain foreclosure rates. The discussion suggests that minority populations, households with less education, less income, and with some kind of financial difficulty are

more prone to foreclosure. However, all the studies did not arrive at the same result. The confounding results, along with their explanations, are discussed below.

- Confounding results

While studying the causes behind high foreclosure rates at the regional level, it has to be taken into consideration that most of the local level studies were done in different regions. So, even though the methods used in the studies were similar, the data sets are unique to each region along with variability in data availability. These dissimilarities might lead to different outcomes. Some of the contradictory outcomes are discussed below.

In a study about residential mortgage default rates in Weber County, Utah, Pedersen and Delgadillo (2007) found that some of the census tracts with less minority and higher income populations seem to be more likely to default on their mortgages. The authors argued that these people overextended themselves in their use of credit and therefore could not keep up with the house payments and other obligations. They also found that some areas with a greater minority population had less mortgage debt and in this case, it was argued that because these people struggle with monthly expenditures even without the mortgage, they have fewer tendencies to borrow at high interest rates.

A study by Calem, Gillen, and Wachter (2004) found a positive correlation between foreclosure rates and percentages of renters in seven large cities in U.S. In other words, higher owner occupancy was associated with lower foreclosure in these areas. But Ong and Pfeiffer (2008) and Edmiston (2009) found a positive relationship between owner occupancy and foreclosure rate in their studies. In another study, Kaplan and Sommers (2009) found that areas in Summit County, Ohio, that had median incomes between $20,000 and $30,000 had the highest rate of foreclosures, but the rate of foreclosure dropped sharply for high income and very low-income categories.

Housing and mortgage characteristics.

Most of the studies discussed in the previous section also incorporated housing and mortgage characteristics to explain foreclosure rates at the regional or local level. While considering housing attributes, Ong and Pfeiffer (2008), Pedersen and Delgadillo (2007), and Bostic and Kwan (2008) found that neighborhoods with a higher proportion of older homes are more likely to have higher foreclosure rates. The reason behind this finding, as given by the authors, is that older homes have more maintenance costs and are often located in neighborhoods with marginalized or vulnerable populations. Additionally, many studies (Ong and Pfeiffer, 2008; Bostic and Kwan, 2008; Immergluck, 2008, 2009) also found a positive association between the magnitude of housing price change and foreclosures. This finding was attributed to the characteristics of 'hot' housing markets that experienced significant increase in housing price during the boom period due to high demand and easy availability of high-risk mortgages. During the downturn when lending dried up and the true nature of the mortgages was exposed, these areas had more overextended households, and thus faced more foreclosures. Housing variables such as vacancy rate and rent to value ratio also showed a positive relationship with foreclosures (Calem, Gillen, and Wachter, 2003, 2004; Ong and Pfeiffer, 2008; and Pedersen and Delgadillo, 2007).

Among all the variables used to explain high rates of foreclosure, the mortgage characteristics have received the most attention. The mortgage characteristics have been identified through different variables such as the prevalence of subprime mortgage, high interest rate home purchase and refinancing of loans, and denial of prime mortgage. In the studies conducted by Pedersen and Delgadillo (2007), Perkins (2009), Calem, Gillen, and Wachter (2004), and Mueller (2006) a positive relationship was found between subprime mortgage or high cost loans and foreclosure rates. The authors argued that minority population, people with low-incomes, people with less education, and older people are most prone to predatory lending due to language barriers, unworthy credit

history, and a lack of understanding of the terms and conditions of the loans. When combined with high unemployment rates, these attributes resulted in high rates of foreclosure.

- Confounding results

The housing and mortgage characteristics have been extensively used to explain foreclosures across neighborhoods. One of the most common outcomes seem to be in areas that experienced large appreciation in housing prices as these areas also suffered the most during the economic downturn. Some authors have found otherwise. These contradictory results are discussed below.

Immergluck and Smith (2005) found that the change in the housing price had a negative correlation with foreclosures in the Chicago area between 1990 and 2000. In support of the previous study, Immergluck (2008) found that in the metro areas where the housing prices escalated, foreclosures are lower even though the level of subprime loans was high. Alternatively, the metro areas with slower housing price increases and high levels of subprime mortgage lending were associated with higher foreclosure rates due to mortgage defaults. The explanation offered for this phenomenon was that homebuyers who had difficulty paying their mortgage in the 'hot' housing market could easily sell or refinance into more affordable mortgages. The initial foreclosure problems in the 'hot' housing market in 2007 was negated by enhancing the loan performance of the subprime lenders and of the securities backed by their loans. But, once the prices in these areas stagnated or declined, the loss rates began to exceed what the securitization models had accounted for and the number of foreclosures increased. In a different study, Ong and Pfeiffer (2008) found that in Los Angeles, areas with more home price appreciation had more foreclosures. Their explanation, which contradicted the previous studies, was that homebuyers moving to the areas with high price appreciation are most

likely to spend beyond their means and thus increase their vulnerability to foreclosure. In the same study, Ong and Pfeiffer (2008) found that in the Los Angeles area, older homes are negatively related to foreclosure rates. However, this result contradicts the findings by Immergluck (2008) as he found that in all the largest metropolitan areas (with more than 50,000 populations) older homes had higher rates of foreclosures.

Spatial characteristics.

There are many studies that looked at the spatial location of foreclosures and high-risk loans in the context of a city or a region. Although these studies do not agree, they can be categorized into three different groups. The first category consists of studies, which concluded that foreclosure and subprime or high-risk mortgages were overrepresented in the suburbs (Mayer and Pence, 2008; Perkins, 2009; and Leinberger, 2008). This conclusion is drawn based on the finding that most of the subprime mortgages are concentrated in rapidly growing suburbs and thus they are likely to have high foreclosure rates. In a study of Contra Costa County in California, Perkins (2009) suggested that high foreclosure rates are concentrated in the outer suburbs that experienced new growth while neighborhoods in the central area have lowest foreclosure rates because the areas around the inner city (in this particular case) have relatively higher median income and include some of the more established and expensive neighborhoods and homes.

The second category consists of studies that found inner-city neighborhoods and the first ring of the suburbs/exurbs to be the areas most vulnerable to the current housing crisis (Garcia, 2003; Immergluck, 2008; Kaplan and Sommers, 2009; Lehnert and Grover, 2008; Newman and Wyly, 2004; Ong and Pfeiffer, 2008). Mayer and Pence (2008) and Perkins (2009) also found similar results. This finding is associated with the finding that inner cities and exurbs have large numbers of minority population, lower

income households, and aging housing stock. These attributes combined with the concentration of subprime mortgages resulted in higher foreclosure rates in these areas.

The third category includes a study by Immergluck (2010) who did not find any association between suburban sprawl and foreclosure. The author suggested that although the suburbs saw higher rates of foreclosure in recent years, it is mainly due to the fact that more high-risk loans originated in these areas during the mortgage lending boom and spatial characteristics did not play any role in it. Therefore, based on the findings of previous studies, it can be concluded that there is no clear consensus among authors about the spatial location of foreclosed properties in the current housing crisis. The same can be said about the association between urban sprawl and foreclosed properties.

Gaps in Current Literature

In the previous section, the main causes behind the current foreclosure crisis were discussed. At the national level, the factor that contributed the most towards the crisis is falling housing prices that subsequently resulted in a weak economy. Although most of the authors have implicitly acknowledged the contribution of these factors, they have not explicitly empirically examined their specific contribution with respect to other factors in explaining the foreclosure crisis. To overcome this gap in the literature we will need to operationalize the concepts and develop carefully specified statistical models. Take for example the role of housing submarkets and its relationship to housing price. The literature about housing markets suggests that the housing submarkets can capture the dynamics of local housing markets (Jones and Watkins, 2009). It is argued that the housing submarket as an analytical framework segregates the local housing market based on spatial and structural attributes so that each segment demonstrates significant housing price difference from other segments. These housing market segments are tagged as submarkets. There are various ways to delineate housing submarkets in a

region including hedonic modeling and clustering techniques. Such methods are suitable for smaller scale geographical areas as one can count the number of clusters formed by contiguous neighborhoods. But for larger study areas, a more sophisticated method, such as max-p regionalization method, is more suitable. This method endogenously clusters areas based on similar housing attributes and contiguity.

The second most important factor behind the foreclosure crisis was the weak economy. The impact of the crisis was felt throughout the nation, but more so in the regions marked by weaker or vulnerable economic preconditions. None of the empirical studies included this feature to distinguish between high and low foreclosure rates. Most studies used mortgage and demographic characteristics to explain variations in foreclosure rates across cities or regions. This gap in the literature can be overcome by incorporating the concept of economic resiliency. This concept, based on economic characteristics (employment diversity, industrial shift, employment by firm size, unemployment rate), helps to determine whether a region is capable of withstanding exogenous economic shock. Therefore, differences in foreclosure rates in different regions of the country can be explained based on the level of economic resiliency.

As discussed earlier, the last decade was marked by periods of housing boom and bust. The period of housing market boom was marked by large numbers of new housing construction resulting in further physical expansion of cities and suburbs. Availability of vacant land, tax subsidies, and easy access to mortgages during this period were some of the major factors behind new housing construction in the suburbs. Immergluck (2010) also noted that most of the subprime loans were generated in the suburbs. But after the sharp decline in housing prices in the latter half of the decade, the rates of mortgage defaults and foreclosures rose in these areas.

Many studies have incorporated the concept of urban sprawl while studying housing foreclosures. The variables used in these studies to represent urban sprawl mostly include distance from central business district (CBD) or mean travel time to work.

Though these measures are useful to make some general connections between foreclosures and urban sprawl, they are at best proxies for sprawl. A more comprehensive and objective measure of sprawl is necessary to comment on the role of suburban expansion on housing foreclosures.

Economic Resiliency and its Measures

The concept of resiliency has been studied in many academic arenas. It has been most extensively studied in the field of ecology where, in 1973, Holling first defined the concept of resiliency. With time it has been applied to various other fields including mental health, disaster management, economics, and urban planning. According to the initial definition, resiliency is the ability of systems to absorb changes and still persist (Holling, 1973). Since its inception, various authors have modified the definition of resiliency according to their respective fields. In the context of urban regions, Hill, Wial, and Wolman (2008) defined resiliency as the ability of a region to recover successfully from shocks that either throw its economy off its growth path or have the potential to throw it off its growth path. This definition is accepted and used in this study. Thus it focuses on the concept of economic resiliency in the context of urban areas.

Even though the concept of resiliency is quite broad and universal in nature, some scholars have tried to quantify it to study the performance of regional economies. Some authors have calculated an index to measure resiliency and compare them between regions, while others have used a combination of different measures.

Ekogen (2008) and Advantage West Midlands (2010) are two private organizations that have each created a resiliency index for cities and regions in the UK. Both indices are a combination of different economic, labor market, and social variables. These variables include percentage of workforce in different types of industries (manufacturing, knowledge base, and public administration), mean wages, unemployment rate, education qualification, household income, and job density. These

variables were included to identify whether an increase or decrease in any of these indicators would increase or decrease the economic resiliency. A similar index has been calculated to measure resiliency at the national level by Briguglio et al. (2008). To measure resiliency at this coarse spatial scale, the authors included indicators such as macroeconomic stability (fiscal deficit, debt), microeconomic market efficiency (participation of foreign banks, control on interest rates), good governance (judicial independence, military interference), and social development (human development index).

To measure economic resiliency at the regional level Chapple and Lester (2007) and Hill, Wial, and Wolman (2008) suggest a combination of a few different measures. Chapple and Lester (2007) compared the level and growth rate of population, employment, industrial diversity, educational attainment of the workforce, and demographic characteristics between two years to measure the degree of economic resiliency and commented on the ability of regions to return back to pre- shock growth patterns based on these changes. Hill, Wial, and Wolman (2008) suggested using location quotients, shift-share analysis, employment diversity, and employment growth by firm sizes to measure economic resiliency. On the other hand, a study by Hill et al. (2010) provided another more unambiguous measure of economic resiliency. A region was labeled, "resilient if, within four years of the onset of the downturn, its annual growth rate returns to the eight year growth rate prior to the year the downturn occurred" (Hill et al., pg 7, 2010). Economic downturn was marked when a shock in a particular year resulted in decline of more than two percent of national growth rate (employment or GDP) from its annual growth rate over the previous eight years.

From the above discussion, it is clear that the measures for economic resiliency vary considerably among studies. After careful evaluation of the various measures, this study decided to focus on the measures used by Hill, Wial, and Wolman (2008). The four most important measures selected for this research are economic diversity, employment

growth by firm sizes, economic competitiveness using shift-share analysis for industrial sectors, and unemployment rate.

Economic diversity.

The concept of economic diversity is defined by Malizia and Ke (1993) as the variety of economic activity which reflects differences in economic structure. This concept was developed in economic literature as an important strategy for the national and local economy to withstand external shocks. It is argued by various scholars like Thompson (1965) that economic diversity provides economic stability to regional economies against shocks and uncertainties. The studies by Bahl, Firestine, and Phares (1971), Attaran and Zwick (1987), Smith and Gibson (1988), Malizia and Ke (1993), and Wagner and Deller (1998) found strong linear negative relationship between regional economic diversity and unemployment rate (an indicator of economic instability).

Over the last 60 years, the concept of economic diversity has been studied extensively by regional scientists. The definition and the methods of measuring economic diversity have evolved over these years. Most of the definitions used by previous authors measured diversity at a specific time rendering it a static and positive concept. This study uses the entropy index to measure the economic diversity of a region. This index measures diversity based on the number of different economic sectors in an area and the evenness of distribution of employment across these sectors. The index increases as an area's economy becomes more diverse. This index is calculated as:

$$\text{Entropy Index} = \sum_{i=1}^{n} Xi.\ln(Xi)$$

Here Xi is the i^{th} sector's share of employment in a region in a particular year when the economy is divided into N sectors.

Being one of the important economic concepts, economic diversity has been applied in many fields of study including urban growth and real estate. Studies conducted

by Chinitz (1961), Quigley (1998), and Wagner and Deller (1998) found that areas that are more economically diverse are not only more stable in their growth, but also experience faster economic growth and recovery. They concluded that when an area's economic fortune is tied to few industries; it experiences slow growth and takes time to recover from economic crisis. In the field of real estate, William (1996) and Mueller (1993) found that real estate portfolios perform much better when areas are considered based on industrial diversity rather than geographical location. Thus, based on these results, it can be concluded that when an area is economically diverse it experiences faster growth, is economically stable, and the return in real estate investments are better. When such circumstances are tested during a period of economic and housing crisis, it can be hypothesized that areas with more diverse economy will incur less mortgage default/foreclosure than those that are less diverse. Some past studies have also found an association between foreclosure and economic diversity (Clauretie, 1988; Eichholtz, 1995).

The studies by Clauretie (1988) and Eichholtz (1995) examined the relationship between economic diversity and mortgage default rates. Clauretie (1988) studied the relationship between residential foreclosure rate and local economic diversification for 109 Standard Metropolitan Statistical Areas (*SMSA*) in the U.S. in the early 1980s. He concluded that foreclosure rate can be explained by economic diversity along with other mortgage characteristics. Eichholtz (1995) looked at the similar relationship for 30 regions in the Netherlands from 1983 to 1990. He concluded that the economic stability indices (measuring change in the local economy from a previous year) rather than economic diversity indices perform better in explaining mortgage default risks. Thus it can be noted that the results from these studies that there is little consensus about how economic diversity is related to mortgage foreclosure rates. Furthermore, none of the recent studies have looked into this relationship in the context of the current housing crisis.

Economic Structure and Firm Size.

Apart from the economic diversity and economic stability measures, another component that plays an important role in a region's economy is the firm size variation within business enterprises/firms. Many economists (Kirchhoff and Phillips, 1988; You, 1995; Shaffer, 2002; Perez-Quiros and Timmerman, 2000) have studied the relationship between firm sizes and economic stability as employment is created and destroyed at the firm level. Thus, while considering the economic stability during economic growth and recession, the incorporation of firm dynamics becomes all the more important.

While studying the U.S manufacturing industry, Hall (1987) found that growth rates of enterprises are negatively associated with their size. Therefore, he concluded that small firms grow faster than large firms. Shaffer (2002) found the similar result across most of the industries in the U.S. when he compared data across 700 cities. The growth rates of firms are intrinsically connected with the employment generated in them. This area has also been studied by many authors including Kirchhoff and Phillips (1988), You (1995), and Garmestani, Allen, Mittelstaedt, Stow, and Ward (2006). All the studies found that small firms add jobs at a faster rate than large firms. Kirchhoff and Phillips (1988) and You (1995) also found that small firms play a bigger role in creating new jobs than large firms, especially during the period of economic recession. You (1995) also suggested that small firms, by their very nature, are a highly heterogeneous group. Thus the functional richness due to the presence of small firms supplements the concept of economic diversity in an area, which is discussed in the previous section. These results clearly demonstrate the role of small and large firms in the economic performance of an area.

Economic competitiveness.

A better understanding of the economy and employment base can have many positive outcomes including improved policies for future growth and development,

proactive strategies to dampen the effects of an anticipated shock and, to direct the attention of the government towards struggling industrial sectors. This pursuit has resulted in quantifying the economic performance of a region using techniques like shift-share analysis.

The shift-share method is a measure of the relative components of the gains or losses of an entity compared with a growth norm (Lipnick, 1987). In other words, this technique measures the growth movement (faster or slower) of local industrial sectors with respect to the growth trends of the industries at the national level. This method is composed of three basic elements: national share, industrial mix, and local share or regional shift. Among the three, regional-shift is the most important component. In many studies, only this component is used as it highlights a region's leading and lagging industries. By comparing the growth rate of local industrial sectors with the same sectors at the national level, the region's competitiveness is ascertained. Although it has some inherent limitations, this method has been used to understand the employment growth at the local level because of its simplicity and limited data requirement.

The Concept of Housing Submarket

The theoretical literature that explains the structure and operation of local housing markets can be divided into two schools of thought. The first is the 'institutional' approach while the second is derived from New Urban Economics, which is a part of neo-classical economic literature. The institutional approach was more prevalent until the 1960s to understand and explain the functioning of the local housing markets. Under this analysis, the authors were concerned with the norms, habits, and cultures of economic agents and the market was considered as a social institution. The main concern was the discontinuities in the urban system and the dynamics of neighborhoods. The researchers tried to derive models of the local housing market that emphasized the co-existence of distinct but interrelated market segments/submarkets. They argue that distinct

submarkets exist because of market imperfections which lead to market disequilibrium. The imperfections include mismatch between the supply (slow adjustment due to stock durability, high transaction costs) and the demand (financial and psychological cost of moving, desire to locate near friends, family and workplace, costs associated with collecting and comparing information about competing heterogeneous vacancies) (Jones and Watkins, 2009).

Since the 1960s, there has been a shift in the theoretical framework to understand local housing market. The new approach suggested that housing markets tends toward multiple equilibria and that each local housing market/submarket is a unitary system exhibiting its own equilibrium price (Goodman, 1978). The housing price difference across space is due to the difference in physical and neighborhood qualities. The price difference of similar units in comparable neighborhoods is temporary and is removed by the arbitrage process. Under this theory, authors acknowledge the existence of price difference between market segments and build models of housing price that are based on the assumption of equilibrium within submarkets (Jones and Watkins, 2009).

Both the approaches discussed above have their limitations. The institutional approach is criticized for being under theorized. There is a lack of precision in defining key concepts, including submarkets which result in difficulties in applying the analytical framework in a consistent manner. Due to these limitations, this approach lacked generalizability. The New Urban Economic approach, though it dominates the existing housing study literature, has its critics. It has been criticized due to its failure to adequately explain the spatial distribution of prices in local housing markets. Its criticism also includes the lack of consideration for institutional roles and the dependence on mathematical sophistication (Jones and Watkins, 2009). Even though the two theoretical schools have different approaches to study housing market and their own set of criticisms, it is clear that there are a number of factors that influence local housing price

difference. Due to this fact, it is important to use an analytical framework to study housing markets that includes the concept of housing submarkets.

Delineation of housing submarkets.

Most of the studies related to housing submarkets can be categorized into three groups based on the method used to delineate submarkets. The first category includes studies that delineate submarkets based on structural characteristics. These studies by Allen, Springer, and Waller (1995), Bajic (1985) and Dale-Johnson (1982) argue that there exists some degree of substitutability for certain kinds of housing units based on occupancy and price so that they can be considered as a part of the same submarket. So, to be in the same housing submarkets, dwelling units should demonstrate a substantial degree of substitutability in terms of value, rents, and occupancy. These studies thus identified submarkets based on attributes such as plot size, floor area, and types of housing unit.

The second category, which is based on submarkets delineated through spatial characteristics, includes studies by Straszheim (1975), Ball and Kirwan (1977), Munro (1986), Bourassa, Hamelink, Hoesli, and MacGregor (1999) and Goodman (1981). These authors argue that housing submarkets are created when homebuyers and sellers are restricted from entering into the market of a particular geographic area. These restrictions can come in the form of housing affordability or availability of mortgage finance, historic characteristics like racial bias on the part of buyer, seller, and real estate agents, search cost, and personal preferences. Most of these studies delineated submarkets using a-priori census and administrative boundaries to identify neighborhoods. These a-priori geographical units are then aggregated based on similar housing and demographic characteristics using principal component analysis. Subsequently, cluster analysis is conducted on principal components to determine the most appropriate groupings.

The third category includes studies that delineated submarkets based on nested spatial and housing characteristics. Such forms of delineation identify submarkets based on housing attributes within spatially segmented geographic areas. Authors like Maclennan and Tu (1996) and Adair, Berry, and McGreal (1996) suggest that both structural and spatial attributes separately or interactively generate housing submarkets. In these studies, the authors first segmented the study area into a-priori spatially segmented areas and then within each area, submarkets are identified based on structural characteristics. Thus the latter are the subset of the former.

Attributes of housing submarket.

Among the housing submarket studies reviewed, the underlying common attribute is the existence of submarkets based on housing price difference between neighborhoods. This is reflected in the use of housing sale price or housing values as dependent variables in identifying submarkets by Adair, et al., (1996), Bates (2006), Bourassa et al., (1999), Fletcher et al., (2000), and Wilhelmsson (2004). The independent variables used to describe the features of the submarkets can be divided into four types. These are housing characteristics, neighborhood characteristics, time dimension, and spatial dimension. Under the housing characteristics, authors have used variables such as lot size, floor area, number of bathrooms, number of bedrooms, housing type (detached, attached), age of house, presence or absence of facilities like fireplace, air conditioning, garage, pool and parking (Adair, et al., 1996; Bates, 2006; Bourassa et al., 1999; Chen, et al., 2007; Fletcher et al., 2000; Goodman and Thibodeau, 2007; Jones, et al. 2002). The variables under neighborhood characteristics include percentage of home ownership, unemployment rate, population and housing density, vacancy rate, household or per capita income, travel time to work, education attainment level, poverty rate, and mortgage approval rate (Adair, et al., 1996; Bates, 2006; Bourassa et al., 1999; Tu, 1997; Wilhelmsson, 2004). Under time dimension authors

included variables like property sale periods/years (Chen, et al., 2007) while the spatial dimension includes distance from the CBD (core, inner, middle, outer, and fringe), and a-priori administrative geographic units (cities, zip codes, census tracts, and census block groups) (Adair, et al., 1996; Bates, 2006; Fletcher et al., 2000; Goodman and Thibodeau, 2007).

Housing submarket and mobility.

Several past studies have analyzed the migration/movement patterns of residents in the housing market (Grigsby, 1963; Rothenberg, Galster, Butler, and Pitkin, 1991). When housing submarkets were used as the analytical framework to study housing markets, scholars specifically looked at mobility patterns of residents with respect to those submarkets (Basu and Thibodeau, 1998; Jones, Leishman, and Watkins, 2004 and 2005). There are two major components in these types of studies. The first component is theoretical, based upon the concept of filtering, initially proposed by William Grigsby in 1963. The second component is the empirical evidence about residents' mobility with respect to submarkets.

The literature is replete with empirical studies (Rothenberg, el al., 1991; Basu and Thibodeau, 1998; Jones el al., 2004 and 2005) that have analyzed residents' mobility within a housing market that is divided into submarkets. Rothenberg, el al. (1991) and Jones el al. (2005) have also addressed the concept of filtering in doing so. Filtering is mainly based on the concept of close substitutability. This concept suggests that the housing market of an area is divided into several distinct segments/submarkets based on various attributes including quality, location and physical attributes. The households move among existing and new dwellings in different submarkets according to their willingness to pay and demand profiles. The empirical parts of the studies looked into migration of residents within and between submarkets using housing property transaction data. The authors looked at the number of origin and destination of these

transactions. The submarkets in all these studies were defined by the housing price difference among different a-priori areas. All the studies indicate that when people buy their homes it is more likely to be in the same housing submarket where they originally reside. The many different arguments behind such a phenomenon are similarity in housing price, quality of housing, locational attributes, and neighborhood characteristics, or in other words, close substitutability. Other factors like search costs, involvement of complex legal and transactional services, and requirement of substantial amount of time, effort and money also restrict people to move far away from their original submarket (Galster, 1996).

Urban Sprawl and its Measure

Squires (2002) defines sprawl "as a pattern of urban and metropolitan growth that reflects low-density, automobile-dependent, exclusionary new development on the fringe of settled areas often surrounding a deteriorating city" (Squires, 2002). Downs (1998) argues that "suburban sprawl has been the dominant form of metropolitan-area growth in the United States for the last 50 years" (Downs, 1998). But in the last two decades, it has come to the forefront and dominated the conversation not only in academic circles but also in social and political arenas. This sudden attention to sprawl is mainly caused by the negative consequences that came out of such development patterns. The discussion about the causes and consequences of sprawl is out of the scope of this study. Instead, this review focuses on the previous attempts to objectively measure it and its relationship to the current foreclosure crisis.

Since suburban sprawl has come to the limelight in academic studies, various authors have tried to quantitatively measure it. This attempt has ranged from measuring sprawl at the local, regional and up to the national level (Ewing, Pendall, and Chen, 2002; Frenkeland Ashkenazi, 2008; Song and Knaap, 2004). One such study, conducted by Ewing, Schmid, Killingsworth, Zlot, and Raudenbush (2003) at the national level,

quantified the extent of sprawl for the largest 448 counties by creating a composite index based on several indicators. This study uses a similar methodology and method to measure sprawl previously developed at the county level (Ewing et al., 2003).

Sprawl and foreclosure.

One of the objectives of this paper is to assess whether suburban sprawl had any effect on the current foreclosure crisis. Therefore, in this study, urban sprawl is measured through a composite index using U.S. Census and Census Transportation Planning Package data for 2000. Though the index provides a picture about the extent of urban sprawl in the beginning of the last decade, it is the most recent data available to calculate this index.

Though none of the previous studies on the current foreclosure crisis used an objective measure for sprawl to explain foreclosures, many have looked at the spatial location of foreclosures and high-risk loans in the context of a city or a region. Among all these studies, though it is hard to find a clear consensus about the spatial location of foreclosed properties, they can be divided into three different categories. The first category consists of studies (Mayer and Pence, 2008; Perkins, 2009; and Leinberger, 2008) that suggested that foreclosure and subprime or high-risk mortgages are overrepresented in the suburbs. This conclusion is drawn based on the finding that most of the subprime mortgages are concentrated in rapidly growing suburbs and thus they are likely have high foreclosure rates. The second category consists of studies (Lehnert and Grover, 2008; Ong and Pfeiffer, 2008; Immergluck, 2008; Newman and Wyly, 2004; Kaplan and Sommers, 2009; Garcia, 2003) that found inner-city neighborhoods and the first ring of the suburbs/exurbs to be the most vulnerable to the current housing crisis. Mayer and Pence (2008) and Perkins (2009) also found similar results. This conclusion is based on the related finding that inner-cities and exurbs have large numbers of minority population, lower income households, and aging housing stock. These attributes

combined with the concentration of subprime mortgages results in higher foreclosure rates in the identified areas. The third category includes a study by Immergluck (2010) who did not find any association between suburban sprawl and foreclosure. The author suggests that though the suburbs saw higher rates of foreclosure in recent years, it is mainly due to the fact that more high-risk loans originated in these areas during mortgage lending boom and spatial characteristics did not play a significant role. Thus, the above studies suggest that though there is no clear consensus among authors about the role of suburban sprawl in the current mortgage crisis, it still has relevance in the analyses of the current housing crisis.

CHAPTER 3

METHODOLOGY

Conceptual Framework

There are three main hypotheses that are tested in this research. These are:

1. *Any urban region with high economic resiliency measures can better cope with housing market downturns*. When an area is economically resilient, its workforce is not dependent on a few industrial sectors and there exists a suitable environment for setting up small business enterprises. Under such conditions, areas facing housing market crisis in the form of mortgage defaults or foreclosures are better suited to absorb shock and recover faster.

2. *Suburbs of urban areas that grew faster during the 2000-2006 economic and housing market boom period have higher rates of housing foreclosures*. During the recent housing market boom caused by easy availability of mortgage to buy new homes, new developments came up in the suburbs. The housing market crisis was primarily felt in the suburbs as most of the subprime mortgages occurred in these areas. Thus, this research also hypothesizes that areas with more urban sprawl experienced higher rates of foreclosures.

3. *Regions that have a greater number of housing submarkets will experience more foreclosures.* Housing submarkets are primarily based on housing size, location, attributes, and value, among other features. People tend to move within similar submarkets (Jones, Leishman, and Watkins, 2004; Jones, Leishman, and Watkins, 2005). In a region with a greater number of different submarkets, the option to move within a similar submarket decreases, as these submarkets are smaller and there is less substitutability. The restriction on mobility combined with the falling housing prices during foreclosure crisis, makes a region prone to higher foreclosure rates.

This study evaluates the interactions among foreclosure rates and economic resiliency, urban sprawl, and housing submarkets through the above hypotheses at the county level in U.S.

Research Scope.

The sample size for this research is 636 urban counties in the United States. There were 3141 counties in the U.S. in 2000, excluding the ones in Puerto Rico. Out of these counties, 671 had more than half of their population living in urban areas. This research categorizes these counties as urban. But only 636 counties are considered as some of the explanatory variables did not yield any data for the other 35 urban counties. According to the 2000 U.S census, these 636 urban counties contained more than 66% (approximately 190 million) of the total population. In the last 10 years, even though the population has grown at a rate of 9.7%, the national economy has gone through a cycle of growth and recession. Together with the national economy, housing values across the country also rose and, subsequently, fell. The period between 2000 and 2006 saw appreciation in housing values as the housing price index (HPI) went up by 79.09 points (Department of Housing and Urban Development, 2011). The U.S. housing market also added 6.4 million housing units during this period. Since 2007, the housing price index has dropped 37.14 points and on average 250,000 foreclosures have been initiated every month (Lender Processing Services, 2011). Additionally, the unemployment rate during the latter period jumped from 4.6% to 9.6% exacerbating the foreclosure rate (Bureau of Labor Statistics, 2010).

Data Collection

As noted above, urban U.S. counties are used as the unit of analysis for this research. The studies related to previous or current foreclosure crises have either focused on the neighborhood level, often represented by census tracts, or regional level represented by metropolitan areas. So far, none have looked into the foreclosure crisis at

the national level with counties as the geographical unit of analysis. As counties are administrative entities, they provide appropriate representation of regions in terms of similar political control, rules, regulations and policies. Another advantage of using counties as the unit of analysis is that the U.S. Census Bureau, along with other major secondary data sources, provides detailed data at this level.

Foreclosure data.

The foreclosure data at the county level was obtained from the Federal Reserve Bank of New York website for December 2009.The data was originally collected by a few private mortgage data collection organizations including LoanPerformance, Lender Processing Services, FirstAmericanCoreLogic, and Mortgage Performance. The Federal Reserve Bank of New York compiles the data at the state or county level and makes it available to the public. The data is categorized based on various types of loans or mortgages. These categories are prime, subprime, Alt-A, jumbo loans, Federal Housing Administration (FHA) and Veterans Affairs (VA), Fannie Mae and Freddie Mac mortgages. The data for subprime and Alt-A mortgages are only available at the state level while the other types are available at the county level. Thus the total number of foreclosures and foreclosure rate for each county was calculated from these latter four (Alt-A, jumbo loans, FHA and VA, Fannie Mae and Freddie Mac) categories. Based on these categories, there were 1.2 million (2.4%) foreclosures out of 50.7 million active loans by the end of 2009 in U.S. At the county level, the rate of loans under foreclosure range from 0% to 13% (See Figure 2). The Federal Reserve Bank of New York provides a disclaimer that the mortgage data it provides only covers 50 to 70% of the total number of mortgages in the U.S.

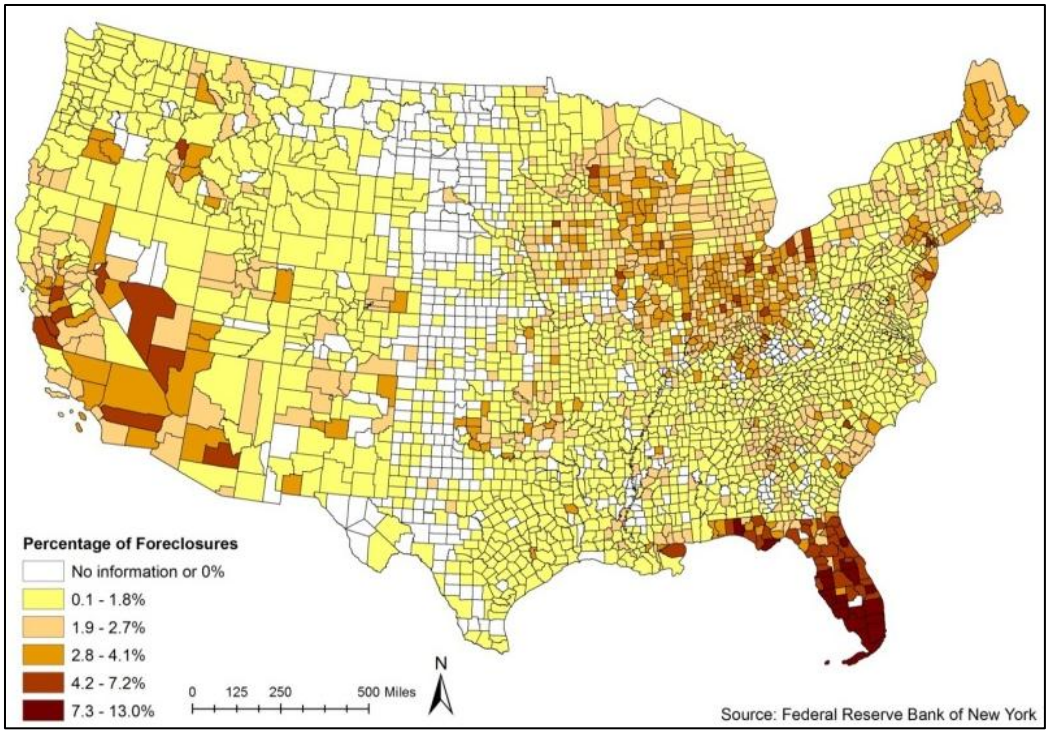

Figure 2. Foreclosure rates in U.S counties (2009)

Explanatory variables: Economic Resiliency, Urban Sprawl, Housing Submarket, and Socio-Economic attributes.

Data used to compile the explanatory variables primarily come from the U.S. Census Bureau. Economic resiliency at the county level was measured using indicators of economic diversity, local industrial competitiveness (local differential shift), growth rate of large and small firms, and the unemployment rate. Urban sprawl is measured using indicators of population density, travel time, and place of work in relation to place to residence. Housing submarket was measured using various housing attributes such as housing value, type, size, occupancy status, and spatial location with respect to place of work. The socio-economic and demographic variables considered in this research are mostly control variables and include population growth rate, percentage of minority population, housing affordability, and mortgage payment characteristics. A complete list of all the explanatory variables can be found in Table 2 along with their sources.

Economic Diversity: Entropy index has been used in this study as the measure of regional economic diversity. Use of this index as a measure of economic diversity was first introduced by Theil (1972). This index measures diversity based on the number of different economic sectors in an area and the evenness of distribution of employment across these sectors. The index increases as an area's economy becomes more diverse. This measure of diversity is popular because it is based on the second law of thermodynamics- the entropy law. The index is relatively easy to compute and requires limited demand for data. Entropy index is calculated as:

$$\text{Entropy Index} = \sum_{i=1}^{n} Xi.\ln(Xi)$$

Where:

Xi = ith sector's share of employment

N = Total number of sectors

The data to calculate this index for all the counties in U.S. were derived from County Business Patterns (CBP) within the U.S. Census Bureau. The data were collected for 2006 as it represents the peak of economic growth in the last business cycle. The data provided employment figures for each of the North American Industry Classification System (NAICS) categories. In this research, only the 21 2-digit NIACS categories were considered. The entropy index of diversity of employment across the 21 industrial categories was calculated for 2006 (See Figure 3) and a higher index score indicates higher level of workforce diversity among these industrial sectors. The purpose of doing so is to test whether economic diversity at the peak of economic growth impacted the foreclosure rate during the subsequent period of economic decline (2007 onwards).

Local differential shift: This is an indicator of economic stability as it measures the growth rate of each industrial sector in each county when compared to the growth of the same industry at the national level. This provides a picture of competitiveness of industries at the county level with respect to national trends. For this research the local

differential shift was calculated for 15 2-digit NIACS industrial sectors between 2000 and 2006 for each county. The 15 industrial sectors considered in this study are based on similarity with previous studies investigating the same issue. The data used to calculate this indicator were obtained from County Business Patterns (CBP) within the U.S. Census Bureau. The formula for the local differential shift is-

$$\text{Local Differential Shift} = LEi(t-1) \cdot \left(\frac{LEi(t)}{LEi(t-1)} - \frac{NEi(t)}{NEi(t-1)} \right)$$

Where:

LE = Local Employment

NE = National Employment

i= i^{th} industrial sector

t = Ending year

t-1 = Starting year

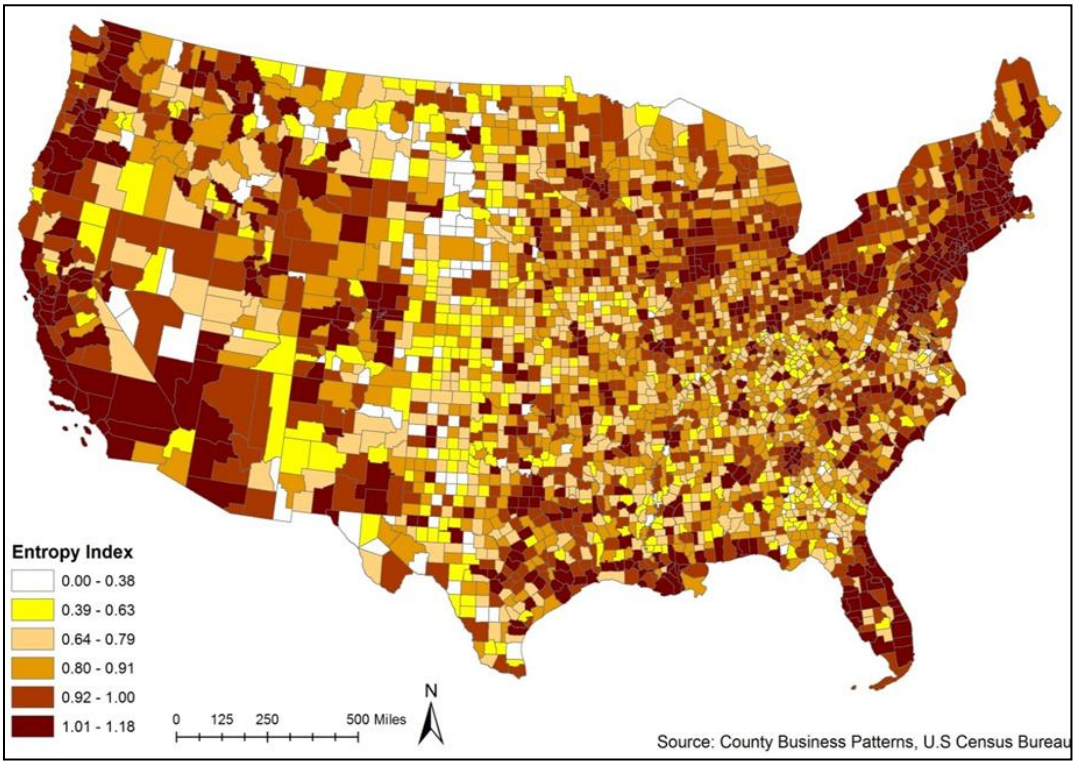

Figure 3. Entropy Index in U.S counties (2006)

Firm size: Business enterprises are an integral part of local and regional economies. They can be broadly categorized as small, mid-size, and large firms. In this research, small firms are considered as those that employ less than 50 people while the large firms are those that employ more than 1000. These threshold figures are based on previous studies and the definitions provided by Organization for Economic Co-operation and Development (OECD). The employment data by different sizes of firms are collected from County Business Patterns for all the years between 2000 and 2008, which is the latest year for which such data is available. This data set provides total number of establishments in each county by their employment range. To calculate the proximate number of employees within these employment ranges, the number of establishments is multiplied by the median value in each range. To get the total employment figure for small firms, the proximate employment for the firms employing 1-49 people are added together. To get the total employment in large firms, the proximate employment for firms employing 1-999 people is subtracted from total employment for each county. Once the employment numbers for small and large firms are computed for all the years, compound annual growth rates for each type of firms for the periods of economic growth (2000-2006) and decline (2007-2008) were calculated.

Unemployment rate: This is also a common variable used in previous studies to gauge economic resiliency of an area. Thus the percentage change in unemployment between 2000 and 2006 and between 2006 and 2009 are considered as measures of regional economic resiliency. The data for these variables come from the Bureau of Labor Statistics.

Suburban Sprawl: Since suburban sprawl has come to the limelight in many academic fields, various authors have tried to objectively measure it. This research follows a methodology similar to that used by Ewing, Schmid, Killingsworth, Zlot, & Raudenbush (2003) to quantify the extent of sprawl. This study calculated the sprawl

index for all the 636 urban counties in U.S. considered for this study using their method but introducing some new variables.

There are six measures that are used to calculate the sprawl index. These six can be categorized in two factors, namely density factor and commute factor. The density factor consists of a) gross population density (persons per square mile); b) percentage of population living in low density (<1,500 persons per square mile) census tracts; and c) percentage of population living in high density (>12,500 persons per square mile) census tracts. These variables were calculated using 2000 U.S. Census data. The commute factor consists of a) percentage of workforce travelling more than 30 minutes to get to work; b) persons travelling outside the county but residing in the same county; and c) persons travelling to the same county but residing outside the county. The first variable was calculated using 2000 U.S. Census data while the last two were calculated using Census Transportation Planning Package data for 2000.

These six variables are combined into one factor using principal component analysis. This single factor, also known as the principal component, represent degree of sprawl within each county while explaining more than half of the variance in the dataset (54.24%). This is also the factor where all the variables load most heavily as seen in Table 1. The principal component derived out of this process has a mean of 0 and a standard deviation of 1. To transform it into the sprawl index it is converted into a scale with a mean of 100 and a standard deviation of 25. This transformation is conducted for the sake of consistency and ease of understanding. The final county sprawl index for the 636 urban counties range from 42 to 352 (see Figure 3). The bigger the value of the index, the more compact the county.

Table 1

Variables used to construct urban sprawl index

Observed Variables	Factor Loading*
Gross population density in persons per square mile	.802
Percentage of population living in densities < 1500 persons per square mile	-.490
Percentage of population living in densities > 12,500 persons per square mile	.809
Percentage of workforce traveling> 30 minutes to go to work	.579
Population that works in the same county but resides somewhere else	.842
Population that resides in the same county but work somewhere else	.821

* Correlation with county sprawl index

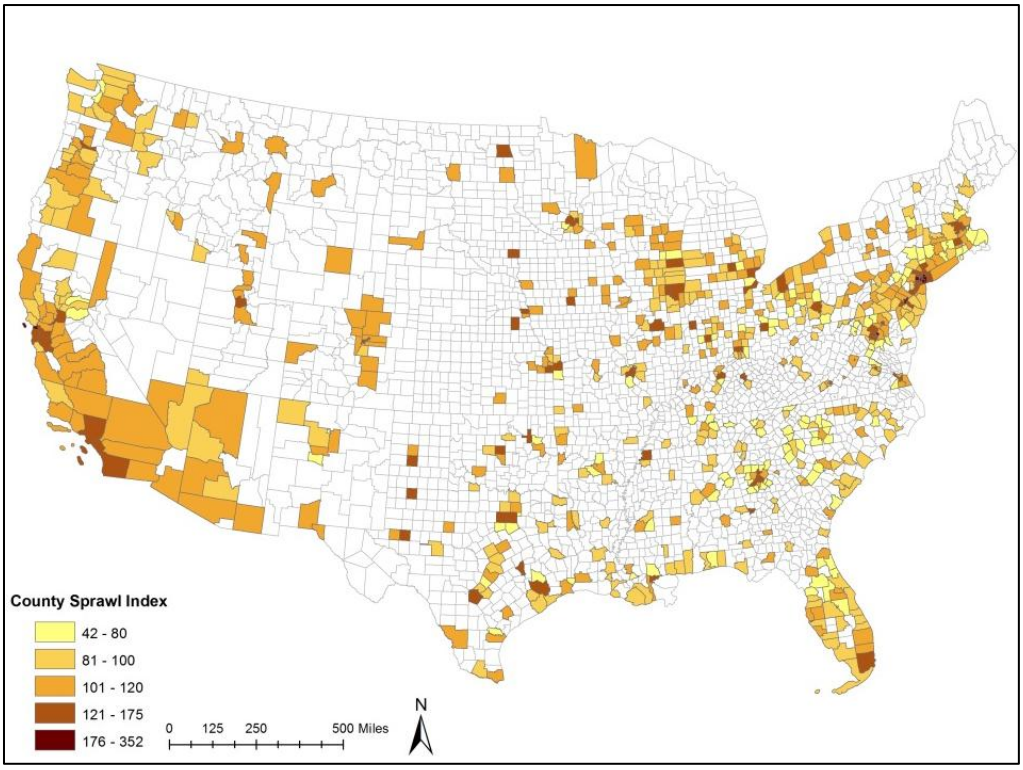

Figure 4. Sprawl index for urban counties (2000); Source: U.S Census Bureau, 2000; Census Transportation Planning Package, 2000.

Housing submarket: There have been numerous studies in the past that looked at the concept of housing submarket. But there is still no consensus among authors about its definition and its measuring techniques. Some of the most common techniques

to delineate housing submarkets in an area are discussed below. This is followed by the discussion of a new spatial clustering method (max-p region problem) and how it is applied to this study.

- Commonly used submarket delineation methods

The majority of the studies dealing with housing economics and housing markets have concluded that housing submarkets exist in an area (Adair, Berry, & McGreal, 1996; Bates, 2006; Bourassa, Hamelink, Hoesli, & MacGregor, 1999; Fletcher, Gallimore, & Mangan, 2000; Goodman & Thibodeau, 2007; Jones, Leishman, & Watkins, 2002; Wilhelmsson, 2004). In spite of the empirical evidence for housing submarkets, there is no standard definition for it. Authors have used different concepts such as market segmentation based on housing tenure, price, dwelling type, type and location of neighborhood (Allen, Springer, and Waller, 1995; Bajic, 1985; and Dale-Johnson, 1982) or economic concept of substitution based on the attributes of the dwelling, and of buyers and sellers (Ball and Kirwan, 1977; Munro, 1986; Bourassa, Hamelink, Hoesli, and MacGregor, 1999; and Goodman, 1981). The later concept varies from the former on the characterization of the spatial dimensions of submarkets. But the recent studies by Maclennan and Tu (1996) have shown that both spatial and structural (housing) characteristics/factors are important in determining submarket dimensions. Based on this finding, housing submarkets can be broadly divided into three categories. These are - structural/ housing attributes, spatial factors, and 'nested' which identifies submarkets based on housing attributes within spatially segmented geographic areas.

A large numbers of studies have identified housing submarkets using the three different techniques mentioned above. In spite of that, the methodological approach across them is more or less the same. Authors have divided their study areas into a-priori geographical units (census block group, census tract, school district, central district, inner

ring, outer ring, and areas based on their geographical alignment to the central district). Once these a-priori areas are identified, authors have used hedonic models to estimate the implicit price of each property attribute. This step is conducted for each potential market segment in order to compare the submarket-specific price for a 'standard' housing unit (Jones, Leishman, and Watkins, 2004). This step is followed by a Chow test to determine if a significant difference exists between the coefficients of the attributes that determine the housing prices in each submarket. The third step involves a weighted standard error test to verify the significant difference between the prices for standard housing units in different submarkets.

- Max-P regionalization method

In this study, a different method (max-p region) is used to delineate housing submarkets. This method is based on the recently developed methods of spatial clustering and is known as regionalization. This method endogenously groups basic units so that intra group homogeneity is maximized while intergroup heterogeneity is also maximized. The units in these groups also have contiguity relationship. In case of max-p region, it aggregates a finite number of geographical areas, *n*, into the maximum number of regions, *p*, so that each region satisfies the following conditions (Duque, Ramos, and Surinach, 2007; Rey, et al., 2011).

1. The areas within a region must be geographically connected.

2. The regional value of a predefined attribute must be greater than or equal to a minimum predefined threshold value. This regional value is obtained by adding up the areal values of the attribute of the areas assigned to each region.

3. Each area must be assigned to one and only one region.

4. Each region must contain at least one area.

The elements that are required for the max-p region problem are listed below (Duque, Anselin, and Rey, 2012):

1. Aggregation variables: These are the set of variables associated with the areas. These variables are used to calculate the dissimilarity between areas as one of the objectives in the max-p region problem is that areas assigned to a region should be similar.

2. Heterogeneity measure: The dissimilarity measured for the areas are combined to calculate the level of heterogeneity for each region and the regional heterogeneity are combined to form a single measure of global heterogeneity.

3. Constrained attribute: This is the attribute variable for which the regional threshold value is set. The regional value has to be more than or equal to this threshold value. This variable does not have to be one of the aggregative variables.

4. Neighborhood structure: This is the information about the spatial contiguity between the areas.

The elements listed above are used in the max-p region algorithm to produce the solution. There are two terms that are calculated in this algorithm based on its objective. The first term controls the number of regions. To begin the process of creating regions, the model assigns areas in the spatial contiguity order zero. These are called 'core areas' or 'seeds' and each region only has one core area. Subsequently, contiguous areas around the core areas are selected to form regions. The second term controls the global heterogeneity of the solution. The measure for global heterogeneity is calculated by summing up all the pairwise dissimilarities between all the areas assigned to the same region. A lower dissimilarity value is sought to achieve homogeneity between areas belonging to the same region (Duque, Anselin, and Rey, 2012). Finally these two terms are merged together to obtain a single value.

To construct the two terms mentioned above, there are certain constraints that are applied by the algorithm. These constraints are- 1) a region cannot have more than

one core area; 2) each area is assigned to only one region with same contiguity order; 3) each region satisfies the regional threshold value requirement; 4) pairwise dissimilarity is considered to calculate the global heterogeneity measure.

Once constrains are applied, a heuristic solution for the max-p region problem is obtained. The solution has two phases- construction phase and local search phase. The construction phase starts with the random selection of an unassigned area called the 'seed' of a growing region. If the attribute value of the selected area is greater than or equal to the regional threshold value, the area becomes a region by itself. Otherwise, one neighboring unassigned area is added to the growing region until the regional threshold value is satisfied. The selection of an unassigned area to a region is based on its effects on global heterogeneity and only if it shares border with at least one already assigned area in a region. Once a growing region satisfies the regional threshold value, a new seed is selected from the set of unassigned areas to start growing a new region. The process of selecting seeds to grow more regions stops when either no area is left to be assigned or when it is not possible to grow a new feasible region from the remaining unassigned areas. These remaining areas are called 'enclaves' and they have to be assigned to one of the already existing feasible regions (Duque, Anselin, and Rey, 2012).

The local search phase improves on the feasible solution by forcing the algorithm to a more intensive inspection of the current feasible solution set. This is done by moving one area from its current region (donor region) to another neighboring region (recipient region) with the hope of discovering a new solution that is better than the current one (Duque, Anselin, and Rey, 2012). If such movement takes place, it needs to follow conditions including- 1) donor region must have at least two areas to allow one area to leave; 2) the removal of an area from the donor region cannot break the spatial contiguity of that donor region; 3) the donor region should still meet the regional threshold value requirement; 4) the area to be moved must share a border with the recipient region.

- Application of max-p regionalization method

The max-p region method joins areas (census tracts) into regions (housing submarkets) for each county based on attribute similarity and spatial connectivity. The aggregate variables based on which the algorithm produces submarkets are listed below. These attribute variables are selected after researching the common traits used in previous studies to demarcate housing submarkets. These five traits represent housing price, size, type, tenure, and geographic location. The data is from the 2000 U.S. Census.

1. Median housing value
2. Median number of rooms
3. Percentage of single family housing
4. Percentage of owner occupied housing
5. Average travel time to work.

The constrained attribute in this case is 'number of housing units'. The threshold value for this attribute is based on median catchment population for school districts in the U.S. This criterion is based on the method used by Goodman and Thibodeau (2003). They used school districts as their a-priori housing submarket units.

The 2000 U.S. Census provides data regarding population served by each of the 14,185 schools districts. They also provide some disclaimers about this dataset as school systems differ for some states. To overcome the issue of overlapping school districts, the elementary school districts are eliminated from the data set as they are most likely to overlap with secondary school districts. The resulting school districts have a median catchment population of approximately 7,000 people. The mean catchment population is not considered because there are many small school districts and few very large school districts, which skews the mean value.

The average number of housing units for school districts is calculated based on average household size in the urban counties considered in this study. The average

household size varies between the range of 2 to 3.75 persons. Based on the these values, number of housing units in the school district catchment area range from 2,000 to 3,500 housing units. The threshold values for the size of housing submarkets are determined at every 500 housing unit intervals within the above range. In other words, submarkets are determined with the threshold values of 2000, 2500, 3000, and 3500 housing units. The max-p algorithm endogenously creates regions (submarkets) based on the five housing attributes mentioned above and these threshold housing unit values. The results obtained from these four threshold values are used in four different regression models to assess the robustness of the submarket variable. Figures 4 and 5 show the spatial arrangement of housing submarkets based on the above criteria for Maricopa County, Arizona and Cook County, Illinois. In these examples, the geographical unit is census tract. So, for each of the counties, contiguous census tracts with the same color represent a housing submarket that has similar housing attributes and meet the threshold value of 3,500 housing units for the constrained variable.

Figure 5. Housing submarkets in Maricopa County, Arizona

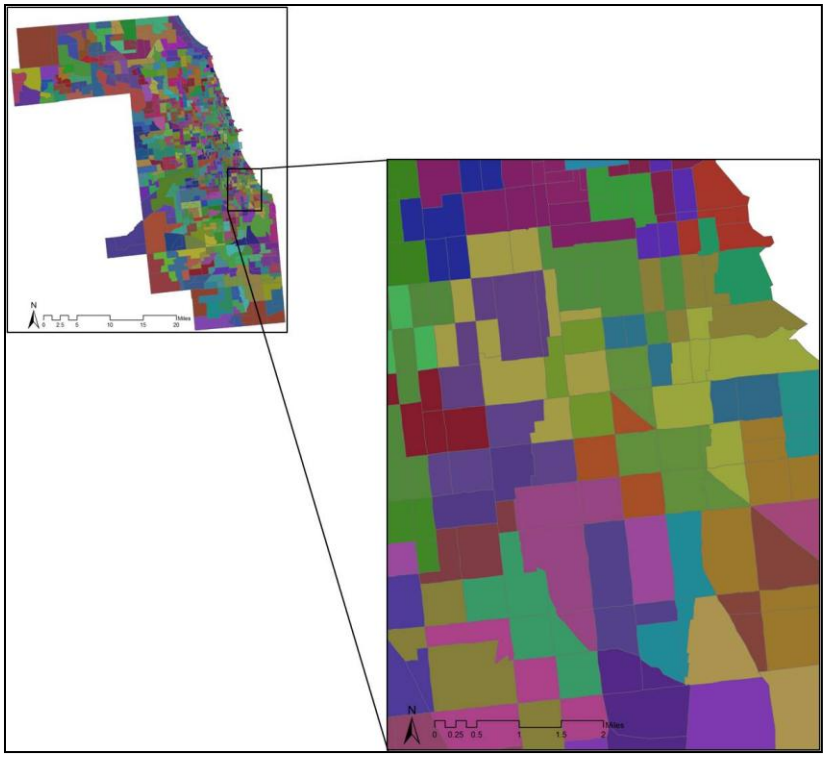

Figure 6. Housing submarkets in Cook County, Illinois

Social, demographic, and economic attributes: The socio-economic variables are selected based on their established association with foreclosure or mortgage default in past studies (Pedersen & Delgadillo, 2007; Perkins, 2009; Mueller, 2006; and Grover et al., 2007). These variables are considered control variables in this study. The purpose of doing so is to identify the relationships between the primary explanatory variables discussed previously and foreclosure rate. The social and demographic variables include percentage change in total population growth rate and percentage change in minority population (non-white population) between 2000 and 2006. Within economic attributes, housing affordability is considered as it captures the change in housing value and median household income. Housing affordability index is calculated for 2000 and 2006. This index is calculated as:

$$\text{Housing affordability index} = \left(\frac{Median\ home\ value}{Median\ household\ income}\right)$$

Another important indicator for housing affordability is included in the form of 'the percentage of the population paying more than 30% of their income on home mortgage in 2006'. Data for all the above variables is collected from the U.S. Census Bureau.

Table 2

Descriptive statistics and data sources for the variables

	Minimum	Maximum	Mean	Std. Deviation	Sources
Dependent variable					
2009 Percent Foreclosure	.002	.130	.020	.017	Federal Reserve Bank of New York
Independent variables					
2006 Entropy Index	.744	1.183	1.058	.070	County Business Patterns (U.S Census Bureau); Authors
2000-06_Compound growth rate_large firms	-.162	.121	.012	.026	County Business Patterns (U.S Census Bureau); Authors
2006-08_Compound growth rate_large firms	-.121	.234	.011	.043	County Business Patterns (U.S Census Bureau); Authors
2000-06_Compound growth rate_small firms	-.071	.093	.016	.018	County Business Patterns (U.S Census Bureau); Authors
2006-08_Compound growth rate_small firms	-.110	.127	.006	.023	County Business Patterns (U.S Census Bureau); Authors
2000-06 Percent population change	-.539	.667	.086	.112	U.S Census Bureau, 2000 Summary File 3; American Community Survey, 2006
2000-06_Perecent minority population change	-1.000	2.450	.200	.388	U.S Census Bureau, 2000 Summary File 3; American Community Survey, 2006
2000_Housing affordability index	1.374	7.678	2.638	.764	U.S Census Bureau, 2000 Summary File 3
2006_Housing affordability index	1.338	13.232	3.941	1.917	American Community Survey, 2006
2000-06_Percent unemployment change	-1.000	1.542	.343	.363	U.S. Bureau of Labor Statistics
2006-09_Percent unemployment change	-.226	3.252	1.047	.471	U.S. Bureau of Labor Statistics
2006_Mortgage > 30% of income	.000	.619	.337	.091	American Community Survey, 2006

Research Procedure

In this research, spatial statistical method is used to analyze the possible connection between foreclosure rates, economic resiliency, suburban sprawl, and housing submarkets. The first step to conduct the analysis includes ordinary least square (OLS) regression followed by spatial regression. These steps are discussed below in detail.

OLS regression.

The first step establishes association between dependent and various explanatory variables while looking for possible spatial autocorrelation in the data. Spatial autocorrelation might exist for a variable if it is correlated with itself in space. It results in spatial clusters. Spatial autocorrelation can be of two types – positive, when high-high or low-low values are clustered together or negative, when high-low and low-high values are clustered together. The positive spatial autocorrelation is a result of Tobler's First Law of Geography which states that areas that are close to each other are more similar than those that are farther away (Tobler, 1970). The presence of spatial autocorrelation in data is a problem and should be resolved because it results in a loss of information and in reduced precision. Global spatial autocorrelation that measures the overall clustering in the data is measured by Moran's I in GeoDa[1]. The value of global Moran's I is dependent on the spatial weights used.

Spatial weights matrix provides a structure of the study area to assess the similarity of location and values among its parts. In this study, queen contiguity-based spatial matrix is used. Under this matrix, all the parts of the study area sharing either a

[1] It is a spatial analysis package developed by Dr. Luc Anselin's Spatial Analysis Laboratory (SAL) in the Department of Geography at the University of Illinois, Urbana-Champaign.

common boundary or point are considered neighbors. While conducting the OLS regression, the spatial weights matrix is incorporated to obtain the value of global Moran's I. Along with the results of the OLS regression and Moran's I value, GeoDa also provides five Lagrange Multiplier (LM) test statistics as one of its sets of the outputs. These statistics later become the guiding tool for choosing particular statistical tests in case statistically significant spatial autocorrelation (Moran's I value) is detected in the data.

Spatial regression.

Once spatial autocorrelation is identified in the data, the second step of the analysis is conducted. This step includes spatial regression as an alternative to OLS regression. As discussed before, GeoDa provides the Moran's I value and five LM test statistics as a diagnostic for spatial dependence. The first two of the LM test statistics include LM-Lag and Robust LM-Lag which are related to spatial lag model. They are followed by LM-error and Robust LM-error which are related to spatial error model. The final one is LM-SARMA which is a higher order alternative, but has little practical use. The selection of one of the two alternative models is guided by a decision tree whose steps are discussed below.

1. If none of the LM-lag and LM-error statistics are significant, OLS provides the best model.

2. If one of the statistics is significant, it should be pursued as it will provide the best alternative.

3. If both the statistics are significant, the robust values should be considered for decision making.

4. If one of the robust statistics is significant, it should be pursued as it will provide the best alternative.

Once the appropriate alternative model is selected and tested, GeoDa provides values for a few more tests in its output to evaluate the alternative model and to detect if some problems still exist in the data. One of these tests is the Breush-Pagan test for heteroskedasticity. It is followed by three classic tests. These are Wald test (W), Likelihood Ratio test (LR), and LM-lag or error test (LM) from the OLS regression. The values for these three tests should be in a certain order (W > LR > LM) to comment on the correctness of the alternative model. The value for the Wald test is obtained by squaring the z-value of the asymptotic t-test.

Spatial regression model.

The data analysis process starting from the construction of the variables to OLS to spatial regression is discussed above. The methodology process followed in this study is explained through Figure 7. Following the sequential steps, this research constructed a spatial error model. The final model is explained below.

$$\text{Ln (PFR)} = \text{K} + \sum_{k=1}^{n} aiREi + \sum_{k=1}^{n} biCONi + \sum_{k=1}^{n} ciSIi + \sum_{k=1}^{n} diHSi + \sum_{k=1}^{n} fiLAMBDAi + e$$

Where:

PFR= Percent foreclosure

RE= Regional economic resiliency indicators

CON= Control variables

SI= Sprawl Index

HS= Housing Submarkets

LAMBDA= Spatially autoregressive error term

K= Constant term

e= Error term

Expected Results

It is expected that the counties that have a diverse workforce will have lower rates of housing foreclosure. This means that regions where the economy is not dependent on few large industries, that encourages growth and development of many small firms, and have lower unemployment rates are better able to absorb the shocks in the housing market. It is also expected that certain industrial sectors like construction, real estate and finance that are associated with the housing market will contribute towards higher rates of foreclosure. On the other hand, based on prior literature, it can be expected that the service industries like heath care and arts and entertainment will stabilize the economy. Based on the growth patterns in the last decade, especially during the period of housing market boom and origination of subprime mortgages, it is expected that suburbs will have a greater number of foreclosures. Based on the dynamics of the housing submarkets in a region, it can also be expected that counties with more housing submarkets will have higher rates of foreclosures.

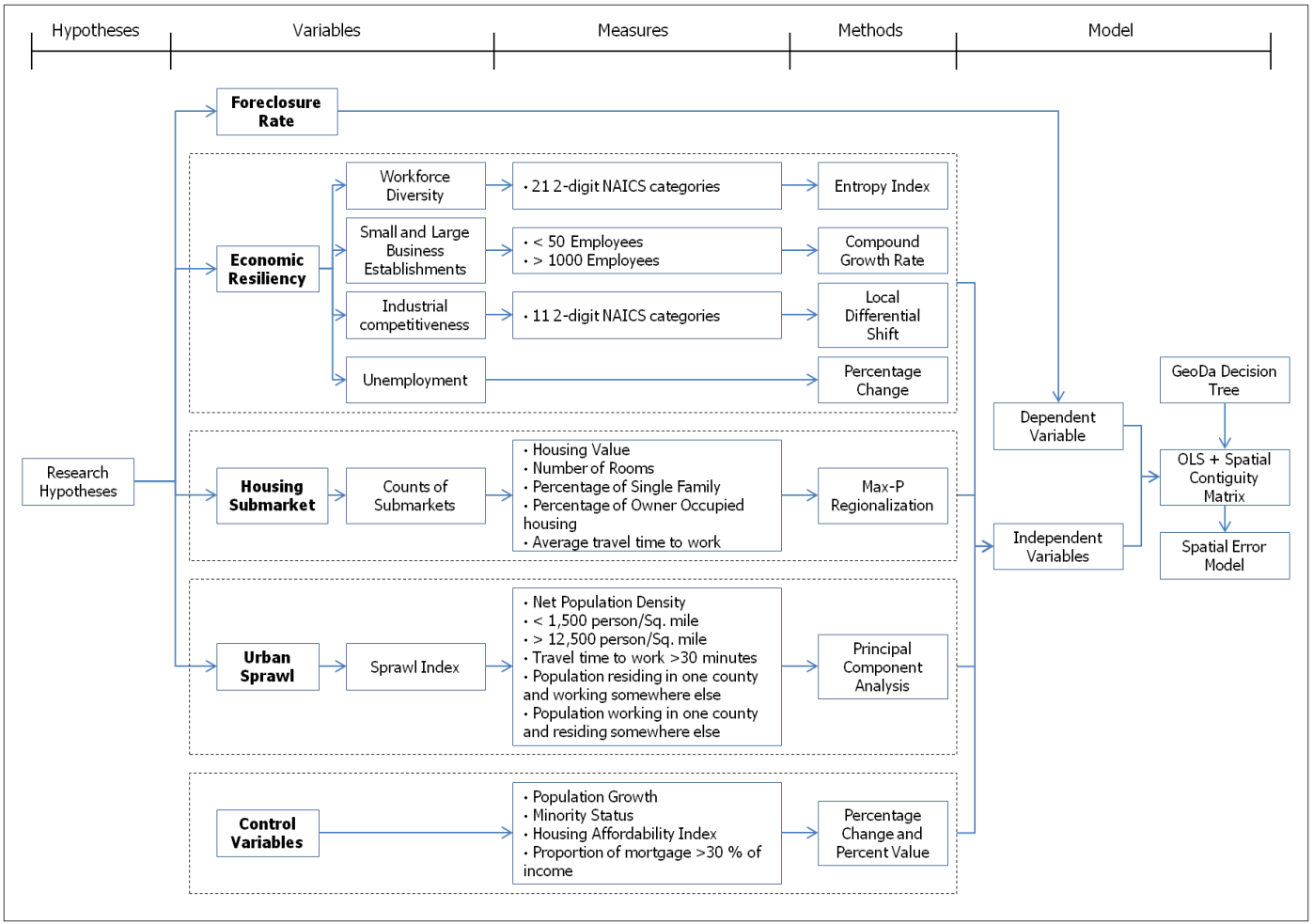

Figure 7. Research methodology flow diagram

CHAPTER 4

RESULTS AND ANALYSIS

Results

This chapter discusses the results for the OLS and spatial error model used in this research. The first part focuses on the diagnostics of the OLS model and identifies its shortcomings. This is followed by the selection of an appropriate spatial model (in this case, spatial error model) as an alternative to OLS. Then, various measures are compared between the OLS and spatial error models to confirm the better fit of the latter. Finally, the diagnostics for the spatial model are analyzed to identify if some problems still remain in the model.

OLS regression with diagnostics

The methodology chapter discussed the various outputs obtained from OLS regression conducted in GeoDa. These outputs include diagnostics for multicollinearity, normality, and heteroskedasticity to indicate any possible problems present in the OLS model. The outputs also include the measure for spatial autocorrelation- Moran's I value along with Lagrange Multiplier (LM) test statistics, which are used to select the appropriate spatial regression model. This section analyzes these outputs for all of the four cases (four different housing submarket sizes) of OLS models. These four models include all the common explanatory variables along with the four different housing submarket sizes. The results are discussed below.

Table 3 lists the diagnostics for multicollinearity, normality, and heteroskedasticity. The multicollinearity condition number is a diagnostic to suggest problems with the stability of the regression results due to multicollinearity. Any number above 30 suggests a problem with the model. As this number for all the cases are above 70, it suggests that the explanatory variables are too correlated and thus provide insufficient separate information.

The Jarque-Bera test for normality of errors suggests that none of the four cases violate the assumption of normality. For a finite sample inference, normality is essential and the current models achieve it. The next three diagnostics are related to detecting heteroskedasticity, i.e., non-constant error variance. The first two test statistics, the Breusch-Pagan and Koenker-Bassett tests, are implemented as tests on random coefficients, which assume a specific functional (*squares*) form for heteroskedasticity. The statistics for these two tests indicate serious problems with heteroskedasticity in all the models. Alternatively, the White statistic does not assume a specific function for heteroskedasticity. This test suggests that heteroskedasticity is not a problem in the models. Such discrepancy can occur when the random coefficient assumption in Breusch-Pagan and Koenker-Bassett tests is not appropriate.

Table 4 lists the last set of OLS model diagnostics to verify whether spatial autocorrelation exists in the dataset. There are six test statistics in this set of model diagnostics that are listed for all the cases. These tests are computed based on the Queen Contiguity weights matrix. Moran's I is the first statistic for which a z-value and associated p-value are provided. The z-values for all the models are very similar (10.9) and the Moran statistic is highly significant. This suggests a problem with spatial autocorrelation, which in this case is due to a mismatch between the scale at which foreclosure occurs and how it is measured by various data sources. As Moran's I is not helpful in suggesting any alternative specification, Lagrange Multiplier test statistics are used. The methodology chapter discussed the steps regarding the selection of the appropriate LM-test statistic. As both the standard LM-Lag and LM-Error statistics are significant, their robust versions are considered. For all the models, the comparison of the significance value for the statistic between the Robust LM-Lag (around $p = 0.35$) and Robust LM-Error ($p < 0.000$) clearly indicates that LM-Error model is the most appropriate alternative to OLS model.

Table 3

OLS Regression Diagnostics

	1			2			3			4		
Multicollinearity Condition Number	73.422			73.287			73.460			73.343		
Test on normality of errors												
Test	DF	VALUE	PROB	DF	VALUE	PROB	DF	VALUE	PROB	DF	VALUE	PROB
Jarque-Bera	2	1.019	0.601	2	0.957	0.620	2	1.025	0.599	2	1.006	0.605
Diagnostics for Heteroskedasticity Random Coefficients												
Test	DF	VALUE	PROB	DF	VALUE	PROB	DF	VALUE	PROB	DF	VALUE	PROB
Breusch-Pagan test	25	49.168	0.003	25	49.137	0.003	25	49.386	0.003	25	49.231	0.003
Koenker-Bassett test	25	45.433	0.007	25	45.493	0.007	25	45.585	0.007	25	45.470	0.007
Specification Robust Test												
Test	DF	VALUE	PROB	DF	VALUE	PROB	DF	VALUE	PROB	DF	VALUE	PROB
White	350	366.732	0.259	350	367.446	0.250	350	368.259	0.241	350	366.760	0.258

1: 2,000 Housing Units Submarket; 2: 2,500 Housing Units Submarket; 3: 3,000 Housing Units Submarket; 4: 3,500 Housing Units Submarket

Table 4

Diagnostics for Spatial Dependence in OLS Models

For weight matrix : 2k_35k_submarket_Queen.gal (row-standardized weights)												
	1			2			3			4		
Test	MI/DF	VALUE	PROB	MI/DF	VALUE	PROB	MI/DF	VALUE	PROB	MI/DF	VALUE	PROB
Moran's I (error)	0.375	10.984	0.000	0.374	10.963	0.000	0.374	10.976	0.000	0.374	10.975	0.000
Lagrange Multiplier (lag)	1	3.481	0.062	1	3.541	0.060	1	3.497	0.061	1	3.488	0.062
Robust LM (lag)	1	0.855	0.355	1	0.811	0.368	1	0.843	0.359	1	0.843	0.359
Lagrange Multiplier (error)	1	113.796	0.000	1	113.351	0.000	1	113.614	0.000	1	113.600	0.000
Robust LM (error)	1	111.171	0.000	1	110.621	0.000	1	110.961	0.000	1	110.955	0.000
Lagrange Multiplier (SARMA)	2	114.651	0.000	2	114.162	0.000	2	114.457	0.000	2	114.443	0.000

1: 2,000 Housing Units Submarket; 2: 2,500 Housing Units Submarket; 3: 3,000 Housing Units Submarket; 4: 3,500 Housing Units Submarket

Maximum likelihood error estimation with diagnostics

Based on the above discussion, it is clear that all the OLS models have multicollinearity, heteroskedasticity, and spatial dependence problems. The LM comparisons indicated that the LM-Error model is a better alternative to OLS. This section discusses whether the spatial error model truly provides a better fit. This process starts with the analysis of the various measures of fit such as Log-Likelihood, Akaike info criterion (AIC), and Schwarz criterion (SC). These measures are provided by the GeoDa software package for both the OLS and spatial error regression models. The R^2 value for the spatial error model is not provided as it is a pseudo- R^2 and thus is not directly comparable with the measure given for OLS results (GeoDa Workbook, 2010). When the values in Table 3 are compared, it can be noticed that value for Log-Likelihood for the spatial error model increased in all cases indicating that it is a better fit. The better fit in the spatial error model is achieved by adding the spatially lagged dependent variable (Lambda) as noticed in Tables 4, 5, 6, and 7. To compensate for the added variable, the values for AIC and SC decrease relative to OLS. All these variations in the measures of fit suggest an improvement in fit for the spatial error specification.

Table 5

Measures of fit for Ordinary Least Square (OLS) and Spatial Error Model (SEM)

	1_OLS	**1_SEM**	**2_OLS**	**2_SEM**
Adjusted R-squared	0.501		0.5018	
Log likelihood	192.171	271.482	192.484	271.507
Akaike info criterion	-332.342	-490.964	-332.967	-491.015
Schwarz criterion	-216.507	-375.129	-217.132	-375.179
	3_OLS	**3_SEM**	**4_OLS**	**4_SEM**
Adjusted R-squared	0.5014		0.5016	
Log likelihood	192.215	271.370	192.328	271.478
Akaike info criterion	-332.429	-490.741	-332.656	-490.958
Schwarz criterion	-216.594	-374.906	-216.821	-375.123

The spatial autoregressive coefficient (Lambda) is estimated around 0.53 and it is highly significant ($p < .0000$) for all the cases (see Tables 6, 7, 8, and 9). In the OLS models, 10 out of 25 explanatory variables are not significant at $p < 0.1$ level. This holds true for all of the four cases. The non-significant variables are: entropy index, percentage of minority population, growth rate of large firm between 2006-08, local shift for manufacturing, finance and insurance, professional and scientific, education, accommodation and food, and public administration industrial sectors, and sprawl index. In the case of spatial error model, there are some changes to the list of above mentioned non-significant variables. The variables that are added to the list are- growth rate of large firms during 2000-06, local shift for retail, information and finance, and health care sector. The variable that became significant at $p < 0.1$ level in spatial error models is the 2006 entropy index. This makes 12 out of 25 explanatory variables significant in the spatial error model. The coefficient values for these 12 significant variables also changed between OLS and spatial error models. The variables whose coefficient value decreased are: entropy index, percentage of population change, housing affordability index for 2000, unemployment rate between 2006 and 2009, percentage of people paying more 30 percent of their income towards home mortgage, local shift for construction industry, and housing submarkets. The variables whose coefficient value increased are- housing affordability index for 2006, unemployment rate between 2000-06, growth rate of small firms between 2000-06 and 2006-08, and local shift for arts and entertainment industry.

Table 6

OLS and SEM results (2,000 HU submarket)

	Ordinary Least Square (OLS): 1				Spatial Error Model (SEM): 1			
Variable	Coefficient	Std.Error	t-Statistic	Probability	Coefficient	Std.Error	z-value	Probability
Constant	-1.91858	0.128	-14.950	0.000	-1.74701	0.107	-16.293	0.000
2006 Entropy index	-0.04119	0.118	-0.348	0.728	-0.16258	0.095	-1.713	0.087
2000-06_compound growth rate_large firms	-0.73984	0.405	-1.825	0.069	-0.10654	0.307	-0.348	0.728
2006-08_compound growth rate_large firms	-0.12898	0.214	-0.604	0.546	-0.03846	0.164	-0.235	0.814
2000-06_compound growth rate_small firms	-4.01199	0.875	-4.583	0.000	-2.55510	0.707	-3.612	0.000
2006-08_compound growth rate_small firms	-2.81987	0.443	-6.365	0.000	-1.92661	0.346	-5.565	0.000
2000-06_percent unemployment change	0.06623	0.023	2.828	0.005	0.07632	0.027	2.848	0.004
2006-09_percent unemployment change	0.16221	0.020	7.930	0.000	0.15607	0.022	7.005	0.000
2000-06_differential shift_Construction	0.00002	0.000	5.561	0.000	0.00001	0.000	4.026	0.000
2000-06_differential shift_Manufacturing	0.00000	0.000	0.744	0.457	0.00000	0.000	0.234	0.815
2000-06_differential shift_Retail	-0.00001	0.000	-2.231	0.026	-0.00001	0.000	-1.605	0.108
2000-06_differential shift_Information	0.00001	0.000	1.841	0.066	0.00000	0.000	0.538	0.591
2000-06_differential shift_Finance and insurance	0.00000	0.000	-0.093	0.926	0.00000	0.000	-0.330	0.742
2000-06_differential shift_Proffesional and scientific	-0.00001	0.000	-1.601	0.110	0.00000	0.000	-1.063	0.288
2000-06_differential shift_Education	-0.00001	0.000	-0.550	0.582	-0.00001	0.000	-0.878	0.380
2000-06_differential shift_Health care	-0.00001	0.000	-1.747	0.081	0.00000	0.000	-0.697	0.486
2000-06_differential shift_Arts and entertainment	-0.00004	0.000	-3.216	0.001	-0.00003	0.000	-2.717	0.007
2000-06_differential shift_Accomodation and food	0.00000	0.000	-0.417	0.677	0.00000	0.000	-0.111	0.912
2000-06_differential shift_Public administration	0.00000	0.000	0.059	0.953	0.00000	0.000	0.001	0.999
Sprawl index	-0.00043	0.000	-0.937	0.349	0.00008	0.000	0.201	0.840
2,000 Housing Units submarket	0.00043	0.000	2.001	0.046	0.00032	0.000	1.870	0.061
2000-06 percent population change	0.59989	0.144	4.157	0.000	0.33054	0.122	2.709	0.007
2000-06_perecnt minority population change	0.01666	0.023	0.712	0.477	-0.01353	0.019	-0.695	0.487
2000_housing affordability index	-0.17863	0.025	-7.109	0.000	-0.22675	0.025	-8.997	0.000
2006_housing affordability index	0.05441	0.011	4.767	0.000	0.07518	0.012	6.161	0.000
2006_mortgage > 30% of income	0.90138	0.127	7.085	0.000	0.68756	0.110	6.251	0.000
Lambda					0.53223	0.030	17.495	0.000

Table 7

OLS and SEM results (2,500 HU submarket)

	Ordinary Least Square (OLS): 2				Spatial Error Model (SEM): 2			
Variable	Coefficient	Std.Error	t-Statistic	Probability	Coefficient	Std.Error	z-value	Probability
Constant	-1.91191	0.128	-14.926	0.000	-1.74727	0.107	-16.317	0.000
2006 Entropy index	-0.04574	0.118	-0.388	0.699	-0.16225	0.095	-1.711	0.087
2000-06_compound growth rate_large firms	-0.73919	0.405	-1.824	0.069	-0.10764	0.307	-0.351	0.726
2006-08_compound growth rate_large firms	-0.13031	0.214	-0.610	0.542	-0.04071	0.164	-0.248	0.804
2000-06_compound growth rate_small firms	-4.01752	0.875	-4.592	0.000	-2.56039	0.707	-3.619	0.000
2006-08_compound growth rate_small firms	-2.82271	0.443	-6.375	0.000	-1.92704	0.346	-5.565	0.000
2000-06_percent unemployment change	0.06613	0.023	2.825	0.005	0.07629	0.027	2.847	0.004
2006-09_percent unemployment change	0.16200	0.020	7.923	0.000	0.15588	0.022	6.995	0.000
2000-06_differential shift_Construction	0.00002	0.000	5.563	0.000	0.00001	0.000	4.041	0.000
2000-06_differential shift_Manufacturing	0.00000	0.000	0.743	0.458	0.00000	0.000	0.233	0.816
2000-06_differential shift_Retail	-0.00001	0.000	-2.245	0.025	-0.00001	0.000	-1.613	0.107
2000-06_differential shift_Information	0.00001	0.000	1.873	0.062	0.00000	0.000	0.546	0.585
2000-06_differential shift_Finance and insurance	0.00000	0.000	-0.087	0.930	0.00000	0.000	-0.317	0.751
2000-06_differential shift_Proffesional and scientific	-0.00001	0.000	-1.549	0.122	0.00000	0.000	-1.057	0.291
2000-06_differential shift_Education	-0.00001	0.000	-0.575	0.566	-0.00001	0.000	-0.888	0.375
2000-06_differential shift_Health care	-0.00001	0.000	-1.708	0.088	0.00000	0.000	-0.690	0.490
2000-06_differential shift_Arts and entertainment	-0.00004	0.000	-3.228	0.001	-0.00003	0.000	-2.711	0.007
2000-06_differential shift_Accomodation and food	0.00000	0.000	-0.447	0.655	0.00000	0.000	-0.120	0.905
2000-06_differential shift_Public administration	0.00000	0.000	0.051	0.959	0.00000	0.000	0.011	0.991
Sprawl index	-0.00046	0.000	-0.992	0.322	0.00008	0.000	0.201	0.840
2,500 Housing Units submarket	0.00055	0.000	2.146	0.032	0.00038	0.000	1.884	0.060
2000-06 percent population change	0.60100	0.144	4.167	0.000	0.33138	0.122	2.715	0.007
2000-06_perecnt minority population change	0.01655	0.023	0.707	0.480	-0.01348	0.019	-0.692	0.489
2000_housing affordability index	-0.17800	0.025	-7.087	0.000	-0.22661	0.025	-8.991	0.000
2006_housing affordability index	0.05425	0.011	4.757	0.000	0.07511	0.012	6.155	0.000
2006_mortgage > 30% of income	0.89821	0.127	7.062	0.000	0.68816	0.110	6.257	0.000
Lambda					0.53174	0.030	17.466	0.000

Table 8

OLS and SEM results (3,000 HU submarket)

	Ordinary Least Square (OLS): 3				Spatial Error Model (SEM): 3			
Variable	Coefficient	Std.Error	t-Statistic	Probability	Coefficient	Std.Error	z-value	Probability
Constant	-1.91691	0.128	-14.925	0.000	-1.74933	0.107	-16.296	0.000
2006 Entropy index	-0.04206	0.118	-0.356	0.722	-0.16099	0.095	-1.695	0.090
2000-06_compound growth rate_large firms	-0.73821	0.405	-1.821	0.069	-0.10697	0.307	-0.349	0.727
2006-08_compound growth rate_large firms	-0.13051	0.214	-0.611	0.541	-0.04001	0.164	-0.244	0.807
2000-06_compound growth rate_small firms	-4.01630	0.875	-4.588	0.000	-2.55863	0.708	-3.616	0.000
2006-08_compound growth rate_small firms	-2.81904	0.443	-6.364	0.000	-1.92570	0.346	-5.561	0.000
2000-06_percent unemployment change	0.06625	0.023	2.829	0.005	0.07623	0.027	2.844	0.004
2006-09_percent unemployment change	0.16203	0.020	7.922	0.000	0.15601	0.022	6.998	0.000
2000-06_differential shift_Construction	0.00002	0.000	5.560	0.000	0.00001	0.000	4.030	0.000
2000-06_differential shift_Manufacturing	0.00000	0.000	0.726	0.468	0.00000	0.000	0.220	0.826
2000-06_differential shift_Retail	-0.00001	0.000	-2.215	0.027	-0.00001	0.000	-1.586	0.113
2000-06_differential shift_Information	0.00001	0.000	1.860	0.063	0.00000	0.000	0.546	0.585
2000-06_differential shift_Finance and insurance	0.00000	0.000	-0.092	0.926	0.00000	0.000	-0.321	0.748
2000-06_differential shift_Proffesional and scientific	-0.00001	0.000	-1.585	0.114	0.00000	0.000	-1.077	0.282
2000-06_differential shift_Education	-0.00001	0.000	-0.563	0.574	-0.00001	0.000	-0.881	0.378
2000-06_differential shift_Health care	-0.00001	0.000	-1.729	0.084	0.00000	0.000	-0.690	0.490
2000-06_differential shift_Arts and entertainment	-0.00004	0.000	-3.220	0.001	-0.00003	0.000	-2.713	0.007
2000-06_differential shift_Accomodation and food	0.00000	0.000	-0.438	0.661	0.00000	0.000	-0.123	0.902
2000-06_differential shift_Public administration	0.00000	0.000	0.055	0.956	0.00000	0.000	0.007	0.994
Sprawl index	-0.00044	0.000	-0.946	0.345	0.00009	0.000	0.224	0.823
3,000 Housing Units submarket	0.00061	0.000	2.022	0.044	0.00044	0.000	1.810	0.070
2000-06 percent population change	0.59936	0.144	4.154	0.000	0.33011	0.122	2.705	0.007
2000-06_perecnt minority population change	0.01660	0.023	0.709	0.479	-0.01342	0.019	-0.689	0.491
2000_housing affordability index	-0.17851	0.025	-7.103	0.000	-0.22679	0.025	-8.995	0.000
2006_housing affordability index	0.05442	0.011	4.769	0.000	0.07520	0.012	6.162	0.000
2006_mortgage > 30% of income	0.89963	0.127	7.066	0.000	0.68812	0.110	6.253	0.000
Lambda					0.53192	0.030	17.476	0.000

Table 9

OLS and SEM results (3,500 HU submarket)

	Ordinary Least Square (OLS): 4				Spatial Error Model (SEM): 4			
Variable	Coefficient	Std.Error	t-Statistic	Probability	Coefficient	Std.Error	z-value	Probability
Constant	-1.91500	0.128	-14.933	0.000	-1.74794	0.107	-16.325	0.000
2006 Entropy index	-0.04319	0.118	-0.366	0.715	-0.16160	0.095	-1.705	0.088
2000-06_compound growth rate_large firms	-0.73929	0.405	-1.824	0.069	-0.10815	0.307	-0.353	0.724
2006-08_compound growth rate_large firms	-0.13024	0.214	-0.610	0.542	-0.04008	0.164	-0.245	0.807
2000-06_compound growth rate_small firms	-4.00923	0.875	-4.581	0.000	-2.55664	0.707	-3.614	0.000
2006-08_compound growth rate_small firms	-2.82021	0.443	-6.368	0.000	-1.92625	0.346	-5.563	0.000
2000-06_percent unemployment change	0.06599	0.023	2.818	0.005	0.07607	0.027	2.838	0.005
2006-09_percent unemployment change	0.16204	0.020	7.923	0.000	0.15601	0.022	7.001	0.000
2000-06_differential shift_Construction	0.00002	0.000	5.555	0.000	0.00001	0.000	4.026	0.000
2000-06_differential shift_Manufacturing	0.00000	0.000	0.723	0.470	0.00000	0.000	0.216	0.829
2000-06_differential shift_Retail	-0.00001	0.000	-2.234	0.026	-0.00001	0.000	-1.602	0.109
2000-06_differential shift_Information	0.00001	0.000	1.859	0.064	0.00000	0.000	0.542	0.588
2000-06_differential shift_Finance and insurance	0.00000	0.000	-0.096	0.924	0.00000	0.000	-0.334	0.738
2000-06_differential shift_Proffesional and scientific	-0.00001	0.000	-1.577	0.115	0.00000	0.000	-1.070	0.285
2000-06_differential shift_Education	-0.00001	0.000	-0.572	0.568	-0.00001	0.000	-0.889	0.374
2000-06_differential shift_Health care	-0.00001	0.000	-1.724	0.085	0.00000	0.000	-0.690	0.490
2000-06_differential shift_Arts and entertainment	-0.00004	0.000	-3.229	0.001	-0.00003	0.000	-2.719	0.007
2000-06_differential shift_Accomodation and food	0.00000	0.000	-0.439	0.661	0.00000	0.000	-0.110	0.912
2000-06_differential shift_Public administration	0.00000	0.000	0.061	0.951	0.00000	0.000	0.014	0.989
Sprawl index	-0.00044	0.000	-0.962	0.336	0.00008	0.000	0.207	0.836
3,500 Housing Units submarket	0.00070	0.000	2.075	0.038	0.00051	0.000	1.869	0.062
2000-06 percent population change	0.59925	0.144	4.154	0.000	0.33047	0.122	2.709	0.007
2000-06_perecnt minority population change	0.01679	0.023	0.717	0.473	-0.01329	0.019	-0.682	0.495
2000_housing affordability index	-0.17846	0.025	-7.107	0.000	-0.22676	0.025	-8.998	0.000
2006_housing affordability index	0.05431	0.011	4.759	0.000	0.07511	0.012	6.155	0.000
2006_mortgage > 30% of income	0.89980	0.127	7.074	0.000	0.68844	0.110	6.260	0.000
Lambda					0.53190	0.030	17.475	0.000

Spatial regression model diagnostics

GeoDa provides two diagnostics for the Maximum Likelihood error estimation. These are the Breusch-Pagan test for heteroskedasticity and the Likelihood Ratio test on the spatial autoregressive coefficient. The values for the Breusch-Pagan test (see Table 8) are not highly significant indicating that the problem of heteroskedasticity has been somewhat addressed. The values for the Likelihood Ratio test (see Table 9) are highly significant suggesting remaining specification problems in the model for all the cases. The Likelihood Ratio test is also part of the three classic specification tests comparing the null model (classic regression specification) to the alternate spatial error model. The other two tests are the Wald test (the square of the z-value for the spatial autoregressive coefficient), and the LM-error test based on OLS residuals. These three classic tests should follow the ordering of: W > LR > LM for finite sample size. The Wald test for all the cases is $17.4^2 = 302.8$ (rounded), the LR test is approximately 158 and the LM-Error test is approximately 113 for all the four cases. It can be noticed that the error models do not violate the expected order and thus indicate a satisfactory model specification.

Table 10

Diagnostics for Heteroskedasticity

TEST	Breusch-Pagan test		
	DF	VALUE	PROB
1	25	37.541	0.0512
2	25	37.655	0.0499
3	25	37.782	0.0485
4	25	37.711	0.0493

Table 11

Diagnostics for Spatial Dependence

TEST	Likelihood Ratio Test		
	DF	VALUE	PROB
1	1	158.622	0.000
2	1	158.047	0.000
3	1	158.311	0.000
4	1	158.301	0.000

Analysis

The analysis of the results is divided based on the various categories of the explanatory variables used in this study. These categories are: regional economic resiliency, urban sprawl, housing submarket, and control variables.

Regional economic resiliency

The explanatory variables included under this category are entropy index, local shift for various industrial sectors, growth rate of small and large firms, and unemployment rate. The explanatory variable of unemployment rate has been used in some of the previous studies to assess regional economic resiliency and some have also used it as a control variable (Ekogen, 2008; Advantage West Midlands, 2010.

One of the important variables considered in this study is the entropy index, which is calculated for all the 21 2-digit NAICS industrial sectors. It is calculated for 2006 to determine its effects during the housing market bust. A higher value of this index indicates a more diverse economy in terms of labor force distribution (in this case among 21 industrial sectors). For all the cases, the coefficient for this variable is negative and significant at $p < 0.1$. For a 1% increase in the entropy index, rates of foreclosure fell by almost 16% in 2009. This result indicates that counties where the workforce was concentrated in a few large industrial sectors during the peak of economic cycle did not fare well during economic recession. This is one of the expected findings for this study. It supports the findings from past studies that concluded that economies whose fortunes

are tied to few large sectors are vulnerable to external shocks. In the case of recent housing market crisis, such economies became weaker and lost more jobs, resulting in large numbers of mortgage defaults by their residents that consequently resulted in higher foreclosure rates. Conversely, counties where the workforce is distributed more evenly among many sectors are better able to handle the external economic stress. In these cases, when a sector is hit by a crisis and loses workers, the whole economy of the region is not impacted adversely and the unemployed workers can join other stable industrial sectors. Thus, these economies are also better able to bounce back to their pre-crisis economic growth pattern. This result therefore supports the hypothesis that workforce diversity makes economies more resilient to outside shocks.

The second set of variables under regional economic resiliency includes the growth rates of small and large establishments/firms during the period of economic boom and bust periods in the last decade. The coefficients for the growth of large firm variables are not statistically significant after controlling for other explanatory variables. This indicates that they did not have any significant impact on mitigating the foreclosure crisis. In the case of small firms, the coefficients for the variables for both the economic boom and bust periods are highly significant ($p < .001$) and negative. This result is complemented by Table 12, which shows that the correlation between growth rates for small firms is positive and highly significant for the boom and the bust periods. In other words, the counties that experienced higher growth rates for small firms during the period of housing market boom also experienced faster growth during the period of market bust. These results suggest that regions where small firms grew faster during the period of economic growth had fewer foreclosures once they were hit by the crisis. Besides that, the same regions were less affected by the foreclosure crisis as they encouraged small firms to grow during economic recession. These results support previous studies that found that small firms are better in helping economies to rebound from crises.

Table 12

Correlation between small firms growth rate for boom and bust periods

		2000-06_compound growth rate_Small firms	2006-08_compound growth rate_Small firms
2000-06_compound growth rate_Small firms	Pearson Correlation	1	.347
	Sig. (2-tailed)		.000
	N	636	636

Out of the 21 2-digit NAICS industrial sectors, the local differential shift is calculated for only 11, based on the categories considered in previous studies. In the spatial error regression model, only 2 sectors are significant. The first is the construction industry, whose coefficient is positive and highly significant ($p < .001$) suggesting that growth in this sector during the period of economic boom resulted in a higher number of foreclosures. This result is expected as regions added more workers in the sectors related to the housing industry, like construction and real estate, during the period of housing market growth. But as the housing prices started to drop (since 2007), the construction industry was hit hard. As the rate of new housing construction dropped, this sector experienced large numbers of layoffs. The regional economies (especially in states like Arizona, California, Florida, Nevada, Michigan, Ohio) that depended on their housing markets as the driver for economic growth were hit hardest as unemployment rose due to layoffs in the housing related industries. The high unemployment combined with restriction on mortgage availability to refinance existing loans resulted in higher rates of residential foreclosures in these areas.

The second industrial sector whose coefficient is statistically significant is arts and entertainment. But in this case, the coefficient is negative suggesting that growth in this sector during the period of economic growth helped regions to mitigate the housing crisis.

This finding is also relevant, as it has been argued by authors such as Richard Florida, that economies that host more jobs for the "creative class" tend to survive and prosper in uncertain times.

The last two variables under regional economic resiliency are unemployment rates during the periods of economic growth (2000-06) and decline (2006-09) in the last decade. These variables are positive and highly significant. This result is expected because when people lose their job, they lose a constant flow of income. In most cases, this results in failure to make mortgage payments and ultimately defaulting and facing foreclosure. As the economic growth slowed down, the unemployment situation got more grim during the period of economic recession. This fact is reflected in the regression result as the absolute value of the coefficient increased and it became highly significant at $p < 0.0000$. The results for all the cases suggest that a 1% increase in unemployment rate during 2006 and 2009 caused the foreclosure rates to increase by more than 15%. Thus, it can be safely concluded that high unemployment does not make a region's economy resilient to withstand economic shocks like the recent foreclosure crisis.

Urban sprawl

The urban sprawl variable is a category in and of itself. This measure was constructed using six separate variables and applying principal component analysis. As discussed before, the sprawl index is a composite index that measures the degree of urban sprawl in every county. A higher value of the index indicates more compact urban built form. This is an important category as many studies have tried to analyze the effects of sprawl on the current foreclosure crisis. Many of these studies have concluded that sprawled areas have more foreclosed properties since these are the areas that added more housing units during the peak of housing growth. But as indicated in the spatial error model, this measure is not significant but has the correct sign. This suggests that, with other variables held constant, regions with more suburban development are

not associated with higher rates of foreclosures. The most plausible explanation for this finding is that foreclosed properties in the suburban areas are a function of high numbers of subprime mortgages disbursed in these areas during the housing boom rather than their spatial location with respect to the central city. Although this finding contradicts conventional thinking, a recent study Immergluck (2010) also arrived at a similar conclusion.

Housing submarket

Similar to urban sprawl, housing submarket is also a category in and of itself. This variable is calculated using a new spatial clustering method called max-p regionalization. This method used six different housing attributes at the census tract level. The census tracts in each county are then joined based upon attribute similarity and spatial contiguity. In the methodology chapter the reasons for selecting four different housing submarkets are discussed. Tables 6, 7, 8 and 9 show the spatial error regression results for these different submarket sizes. All of these tables show that the housing submarket variable is positive and significant for every case. These results suggest that with the increase in the number of housing submarkets in a county, the number of foreclosures also goes up. This result holds true for all four sizes of submarkets, indicating robustness of this variable in relation to housing foreclosures when all other variables are held constant. A plausible explanation for such a finding is that a high number of submarkets increases search costs and decreases the availability of comparable housing, thereby slowing down the market transactions. As foreclosed properties stay in the market longer, they lose their value due to factors like poor maintenance and increase in crime rate and illegal activities. Under such conditions, when residents try to move to a different location within a region, their mobility is restricted if large numbers of housing submarkets exist in a region. This is so because as the number of dissimilar submarkets/segments increases, there is less substitutability in

terms housing types for the residents to choose from. This leads to restriction in intra-urban migration for affected property owners, and possibly higher inter-urban migration, which subsequently leads to more foreclosures in these markets.

Control variables

The various control variables selected for this study are based on findings from previous studies related to the recent housing crisis. These variables (or variables similar to them) were found to have a significant association with foreclosure rate or mortgage default rate. The first measure in this set is the growth rate of the population during the period of economic boom. This variable shows a positive and significant relationship to foreclosure rate. This suggests that the counties that grew faster and added more people during the economic growth, faced hardship during economic recession. This relationship seems to suggest that growing regions were also overbuilt. After the bust of the housing market bubble, a lot of the newly constructed housing units went under water due to the rise in unemployment and difficulty in mortgage refinancing.

The second variable under this category is the change of minority population in the urban counties in the U.S. Traditionally, most authors have argued that the minority populations are more affected by the current foreclosure crisis. In the regression models for this study, though the variable for minority population show the correct sign, it is not statistically significant. This indicates that after controlling for income and housing value, the minority population is not disproportionally affected by the foreclosure crisis. This also means that the current crisis negatively affected people irrespective of their race. This finding is also confirmed by Immergluck (2010) in one of his recent studies. When looking at the relationship between the Hispanic and Black population and real estate owned (REO)/foreclosed homes in the largest 75 MSAs, Immergluck (2010) found that minority populations did not have significantly different levels of REO accumulation.

The next set of variables comes in a pair as they measure housing affordability in two different time periods. The first variable measures housing affordability in 2000 i.e., at the beginning of the housing market boom. The results indicate that after all the other variables are accounted for, housing affordability index in 2000 is negative and highly significant. So, areas with high levels of housing affordability in 2000 had lower levels of foreclosure at the end of the decade. On the other hand, the housing affordability index at peak of housing market bubble (2006) is positive and highly significant. This clearly suggests that during the period of housing market boom between 2000 and 2006, the housing values rose much faster than income levels, making housing less affordable in many urban areas. During this bubble, people could afford bigger and higher value homes because high-cost subprime mortgages were offered by the mortgage lending institutions. After mid-2006, as home prices dropped and people started to lose their jobs, they owed more than their home's worth. This eventually led to foreclosures in large numbers.

The last variable in this category is the percentage of people who paid more than 30% of their income towards their home mortgage in 2006. This variable is also a measure of housing affordability as housing is considered unaffordable if one has to pay more than 30% of their income towards it. Similar to housing affordability index in 2006, this variable is also positive and highly significant. It can be noticed in all the regression models that with a 1% increase in such population, the rate of foreclosures increase by around 69%. This finding reaffirms that during the last housing market cycle, housing values increased disproportionately compared to individual income, resulting in high rates of foreclosures once the housing market crashed.

Summary of results

Based on the analysis discussed in this chapter, this study found that regions that are economically resilient were better able to absorb the shock of the recent housing

crisis. Higher workforce diversity and growth of small business enterprises contributed towards high economic resiliency. Additionally it was found that regions where housing markets are more segmented due to numerous numbers of housing submarkets, experienced higher rates of foreclosures. Based on the association between urban sprawl and foreclosure rates found in the results, it can be said that sprawl did not play a significant role towards the housing crisis. It was also found that decrease in housing affordability during the last decade worsened the impact of the crisis.

CHAPTER 5

CONCLUSION

Conclusions

This chapter summarizes the results of this study and highlights the theoretical contributions made through this research. There are three primary conclusions that can be drawn.

Significant conclusions

- Economic resiliency matters

It was found that economic resiliency of regions played a role in mitigating the latest foreclosure crisis in the U.S. There are various parameters that represent economic resiliency including workforce diversity, growth of firms, growth of industrial sectors, and unemployment rate. Based on the findings discussed in the previous chapter, it can be concluded that counties that have a more diversified workforce are better able to handle the external economic and housing shocks. This is because their economic fortune is not tied to any one or a few large industrial sectors. It was also found that regions that encourage small businesses and where they are an integral part to the economy, bounce back to their pre-crisis growth pattern faster than other regions where small businesses are less dynamic. Another important but expected result was that the rapid growth of the construction industry made the foreclosure crisis worse. This result is understandable because many regional economies were dependent on housing related sectors for their economic growth. With the significant drop in housing prices after 2006, these sectors lost a large number of jobs. Therefore, a high unemployment rate on top of a weakening economy made the foreclosure crisis worse. Conversely, industrial sectors involving more creative occupations like arts and entertainment, performed better in the face of the

crisis. It also helped mitigate the crisis to some extent. Similar to other economic resiliency studies, this research found that higher rates of unemployment, especially during periods of economic bust, lead to deeper crisis. This phenomenon was seen during the second half of the last decade.

- Urban sprawl is not relevant to the foreclosure crisis

Since the latest foreclosure crisis came to the forefront, a large number of studies have looked into the association between foreclosed properties and urban sprawl. But there is no consensus among authors about the type of association, as some have found positive correlation while others did not find any relation. This study concurs with the authors belonging to the latter group. The results indicate that sprawl is not associated with higher rates of residential foreclosures. Higher rates of foreclosures in many suburbs are a reflection of high numbers of subprime mortgages being generated in these areas and not their spatial location. More subprime mortgages originated in the suburbs since these areas were growing at the time when easy access to home finance resulted in more new and larger residential developments. In other words, foreclosures are a function of the type of home loans and not home location.

- Excessive housing submarket fragmentation contributes to the problem

As mentioned above, the recent foreclosure crisis generated considerable scholarly work. But none of the studies looked into the local housing market dynamics while explaining higher rates of foreclosures in various parts of the country. This study fills this gap by analyzing the association between local housing submarkets and foreclosure. Based on the results, it can be concluded that regions with a high number of housing submarkets are more susceptible to foreclosure crisis. This finding can be explained by the fact that too many dissimilar housing submarkets brings down the

degree of substitutability in the housing market. During a period of crisis when intra-urban migration can help one avoid its affects, lower substitutability results in worsening of the crisis.

Other findings

This study also found some important results through the control variables used in the regression models. Most of them confirm similar results found in past studies. Based on these results, it can be concluded that the regions that grew faster during the last decade experienced higher magnitude of foreclosure crisis. Additionally, contrary to popular belief, the minority population is not more likely to face foreclosure, thus suggesting that the latest crisis impacted people across racial and ethnic lines. It has also been shown in previous studies that the housing prices rose sharply during the housing market boom (2000-06). But during the same period, the household income did not see a comparative increase. This resulted in decline in housing affordability. Once the economy slowed down and housing prices dropped, the new homeowners owed more than their home's value to the mortgage lending institutions. The findings of this study corroborate the above theory. The results show that at the peak of the housing market bubble, people paying more than 30% of the income towards their home mortgage were more likely to face foreclosure. Therefore, it is reasonable to conclude that decline in housing affordability played a significant role in worsening the foreclosure crisis.

Contribution of the Study

This research focused on understanding the role played by regional economic resilience, housing market dynamics, and urban spatial patterns in mitigating or exasperating the recent housing crisis at the county level in the U.S. This research provides a theoretical understanding of regional economy in the face of a crisis with the intent to promote a more in-depth understanding of various components of regional economic structure while exploring various dimensions associated with the housing crisis.

Part of this research also looked into local housing market segmentation in the form of housing submarkets to distinguish between areas with contrasting abilities to withstand an external shock. This research also tries to address the theoretical underpinning to the contentious relationship between urban sprawl and higher concentrations of foreclosed properties. This study thus helps in identifying and exploring the various dimensions of the recent housing crisis with an objective to understand their theoretical relationship with regional economy, housing markets, and urban development patterns so that certain conditions can be identified for long term and stable economic growth, increases in housing affordability and access to housing.

The methodological work developed in this study to compute economic resiliency, to delineate housing submarkets, and to measure urban sprawl also distinguishes this research from other existing studies. The regional economic resiliency in the context of U.S. counties is measured for the first time using four different sets of indicators (workforce diversity, growth rate of small and large firms, local differential shift for industrial sectors, and unemployment rate). Furthermore, this study is also among the first studies to objectively measure urban sprawl at the county level and examine its possible association with residential foreclosures. Previously, researchers have either considered pseudo sprawl measures such as mean travel time to work, distance from CBD, or binary dummy variables for areas based on their distance from CBD. In this study, the composite urban sprawl measure was calculated using population density at the census tract level and commute pattern of residents and workers for each county. This study also offers a new methodological approach to delineate housing submarkets within a local housing market. This approach uses a new spatial clustering method (max-p region) that combines contiguous census tracts based on various similar housing attributes. This method of identifying submarkets is different from other commonly used methods as in this case submarkets are endogenously created by the data.

This is also one of the first studies looking into the relationship between local housing market dynamics (measured by housing submarkets) and housing foreclosures.

As a part of the methods used, this study takes into consideration the spatial autocorrelation of foreclosures in its final regression model. This approach is better than previously used OLS models while understanding the probable associations between various explanatory variables and foreclosure rates. This is so because the presence of spatial autocorrelation results in loss of information and less precision. This is also one of the few studies that looks at the current foreclosure crisis at the national level with counties as units of analysis. This regional approach to the crisis contributes toward generalizing the findings for much of the country.

Planning Implications

The flow of capital in the U.S. housing market in the first half of the last decade made housing finance easily available in the U.S. urban areas. But many of the first lien and refinance mortgages disbursed during this period had high risk terms associated with them. In the second half of the last decade however, the flow of capital dried up and the housing values dropped sharply. This, combined with a high unemployment rate, resulted in high rates of mortgage defaults and foreclosures in the U.S. Local governments and municipalities were faced with the effects of the housing crisis even though they did not play a significant role in creating the problem. Under such circumstances, planners have a large role to play, not only to address the exiting crisis, but also to propose policies to make urban areas more robust towards future housing and economic uncertainties. The conclusions drawn from this research can help planners to direct their attention to some areas of concern.

To make urban regions more resilient to future economic shocks, policy makers and planners can help formulate long term economic policies that aim for regionwide workforce diversity. There are various ways to achieve this diversity and one of them is

to encourage small businesses. Cities can work with local banks and entrepreneurs to expedite the process of setting up locally owned small business establishments. Additionally, cities can use strategic marketing and collaborate with the local business community to attract workers and investments for industries that are performing below par. They can also form partnerships with educational institutions to train the local labor force so that workers can more easily adapt to the changing business environment and technological requirements. These steps will ensure that a region's growth is led by a group of industries in which small businesses play a vital role and where workers have adequate skill sets to apply themselves in different sectors. Such programs might include enhancing computer skills, familiarity with handling heavy and sophisticated machinery, and learning managerial skills among workers. These occupational trails among workers are important in a building resilient local economy as these skills are applicable across multiple industrial sectors.

The findings based on the housing submarket variable suggest that households residing in neighborhoods with foreclosed properties are adversely affected due to the continuing decline in housing values. This effect is worsened by limited number of comparable housing options if that region has a large number of dissimilar housing submarkets. Under such conditions, effective planning measures can play a vital role in stopping such neighborhoods from further deterioration. Local neighborhood organizations can work with delinquent homeowners and mortgage lenders to forge an agreement that puts the responsibility on homeowners and lenders for maintaining the vacant foreclosed property to a certain standard. This step will stop the negative spillover effects of foreclosed properties from spreading to the whole neighborhood. Local housing administrative bodies can work with federal government to buy some of the foreclosed properties at appropriate areas for reduced prices and list them under any of the federal housing programs like HOPE VI, Section 8, or other voucher programs. This will stabilize the local housing market and will get the vacant properties reoccupied at a faster rate.

When it comes to providing affordable housing options across a region, planners can advocate for various affordable housing programs such as limiting exclusionary zoning, opposing any policies that discourage renting in affluent communities, and providing different tenure options in most neighborhoods. Many of these suggestions prescribed through this research have also been advocated by other scholars like Dan Immergluck, Jeff Crump, Michael Johnson, David Turcotte, and Felicia Sullivan.

Though this study did not find any association between urban sprawl and foreclosure rate, it concluded that foreclosures in the suburbs are a function of high concentration of subprime mortgages in these areas. Local planners can work towards steps to avoid such housing crisis in the future. They can recommend to the federal government various neighborhoods within a region where local banks can better meet the credit requirements of the community under the Federal Community Reinvestment Act (CRA). They can also learn from the current crisis and be proactive about understanding the trends in the housing mortgage market and rally against future speculative housing developments in their regions. Additionally, planners can provide targeted counseling to vulnerable populations and potential future homeowners so that they don't fall prey to high-risk loan lenders.

Limitations of the Study

Some of the limitations of this study come in the form of two important assumptions considered in this study. One of the assumptions is related to the urban sprawl index. This study assumes that the urban growth patterns experienced in each urban county in the last decade was similar to the patterns experienced through the decade of the 1990s. The reason behind this assumption is the data availability to calculate the sprawl index. The data for the six variables used to calculate the index comes from the 2000 U.S. Census Bureau, which provides a snapshot of the urban development pattern at the starting of the last decade. The foreclosure rates in this

study are measured for 2009. This study analyzes and comments on the possible association between sprawl and foreclosure rates while acknowledging the data limitations.

The second assumption is similar to the first as it also rises from data limitation, but for housing submarkets. This study assumes that the housing attributes like housing type, size, and tenure at the census tract level did not go through significant change during the last decade. But for the attributes like housing value that might have changed, the change was proportional to the housing price in the 1990s for all the census tracts in a county. This assumption is formulated because the six variables used to construct housing submarkets are obtained from 2000 U.S. Census Bureau. This study comments on the association between housing submarket and foreclosure rates in 2009 while acknowledging this disconnect in the data.

Another limitation to this study comes from the limited publicly available foreclosure data. As discussed before, the foreclosure data for this study is collected from NYFRB for December 2009. One of the disclaimers for this data mentions that it does not include subprime and Alt-A loans data and it only contains 50-70% of all the active loans. The literature on the current crisis provides evidence that a high percentage of current foreclosures are the result of defaults on subprime and Alt-A loans. Thus, the absence of data for these types of loans results in failure to capture the full magnitude of the current foreclosure crisis. The foreclosure data used in this study is a result of defaults on home mortgages from FHA, VA, Fannie Mae and Freddie Mac. The subprime mortgage data at various geographic scales are collected by many private organizations. But because of the for-profit nature of these organizations, foreclosure data is only available for a substantial fee. This is one of the reasons why the NYFRB did not include such data in their data set. The mortgage data provided by the NYFRB is compiled from data provided to it by private organizations.

An added limitation to this research is linked to the possible association between housing submarket and urban sprawl. The number of housing submarkets for the urban counties is calculated using Max-P regionalization method. One of the conditions under this method is that each region/submarket should meet a predefined threshold value for an attribute. This determines the size of an individual submarket. Consequently it also determines the total number of housing submarkets in a county. In this study, the predefined threshold value ranged from 2,000 to 3,500 housing units. So, based on the above condition and the size of each submarket, it can be suggested that counties with more population and higher density development will have more numbers of housing submarkets. The sprawl index variable considered in this research also points towards the type of development in every urban county. Higher sprawl index values suggest denser development pattern.

The above discussion indicates a possible overlap between the development trends captured by the housing submarket and sprawl index variables. A bi-variate correlation is conducted between sprawl index and numbers of housing submarkets to verify this overlap. Table 13 indicates that the correlation coefficients between sprawl index and housing submarket variables (for all four submarket sizes) are positive and highly significant. This suggests that any urban county with high density development will also have a higher number of housing submarkets. It can also be said that some of the trends captured by the housing submarket variable are partially explained by the sprawl index variable and vise versa. So while interpreting the results of this study, the above relation needs to be considered.

Table 13

Correlation between sprawl index and housing submarkets

		Sprawl Index	2,000 HU Submarket	2,500 HU Submarket	3,000 HU Submarket	3,500 HU Submarket
Sprawl Index	Pearson Correlation	1	0.505	0.504	0.503	0.501
	Sig. (2-tailed)		0.000	0.000	0.000	0.000
	N	636	636	636	636	636

Future Work and Directions

The recent foreclosure crisis has attracted considerable attention from many scholars. However, the multiple dimensions of this problem, as revealed during this research, suggest that more work still needs to be done to get to the bottom of this crisis. Some of the opportunities for future research are discussed below.

This study analyzes the relationship among various economic, housing, and development patterns and their impact on foreclosure rates. Based on the study results, some generalized conclusions are drawn for the whole country. To understand the impact of the housing crisis in different parts of the country, the future studies can also divide their study area based on various regions (e.g.: Census regions, industrial regions). This will indicate whether unique regional trends exist when it comes to housing crisis and its mitigation.

As discussed under the section of limitation of this study, some of the explanatory variables like the sprawl index and housing submarket are calculated using the 2000 U.S. Census Bureau data. This renders these variables dated in the context of this study. So the future studies can use the detailed data from the most recent 2010 U.S. Census to calculate the explanatory variables and consequently look into their relationship with foreclosure rates. Such studies will be able to comment on the urban growth and housing changes that took place in the last 10 years independently and in

the context of the housing crisis. With these future studies, a longitudinal comparison can be made between the results for the beginning and end of the last decade.

The nature of this study is empirical, the methods used are quantitative in nature, and the study is done at the national level. Future studies can look into the foreclosure problem at a local level using a qualitative approach. Such studies might include analysis of the local housing policies along with interviews of local policy makers, developers, real estate agents, workers, and residents. Such qualitative data can shed light on local level issues, concerns, and shortcomings that vary considerably among regions. The results from these studies can complement the empirical studies so that the foreclosure crisis can be analyzed in both an objective and subjective manner at different geographical scales.

REFERENCES

Adair, A., Berry, J. N., & McGreal, W. S. (1996). Hedonic modeling, housing submarkets and residential valuation. *Journal of Property Research*, 67-83.

Advantage West Midlands. (2010, October 26). *Community Economic Resilience Index.* Birmingham: Advantage West Midlands.

Allen, M. T., Springer, T. M., & Waller, N. G. (1995). Implicit pricing across residential rental submarkets. *The Journal of Real Estate Finance and Economics*, *11*(2), 137–151.

Alonso, W., Technology, M. I. of, & University, H. (1964). *Location and land use*. Harvard University Press.

Attaran, M. (1986). Industrial diversity and economic performance in U.S. areas. *Annals of Regional Science*, *20,* 44-54.

Attaran, M., & Zwick, M. (1987). The effect of industrial diversification on employment and income: A case study. *Quarterly Review of Economics and Business*, *27*, 38-54.

Ball, M. J., & Kirwan, R. M. (1977). Accessibility and supply constraints in the urban housing market. *Urban studies*, *14*(1), 11.

Bahl, R. W., Robert F., and Donald P. (1971). Industrial diversity in urban areas: Alternative measures and inter-metropolitan comparisons. *Economic Geography*, *47*, 414-25.

Bajic, V. (1985). Housing-Market Segmentation and Demand for Housing Attributes: Some Empirical Findings. *Real Estate Economics*, *13*(1), 58–75.

Baldwin, J. R., & Brown, W. M. (2004). Regional manufacturing employment volatility in Canada: The effects of specialization and trade. *Papers in Regional Science*, *83*(3), 519–541.

Bostic, R. W., & Kwan, L. O. (2008). Mortgage, Risk, Homeownership among Low and Moderate Income Families. *American Economic Review*, 310–314.

Bourassa, S. C., Hamelink, F., Hoesli, M., & MacGregor, B. D. (1999). Defining Housing Submarkets. *Journal of Housing Economics*, 160-183.

Briguglio, L., Cordina, G., Farrugia, N., and Vella, S. (2006). Conceptualising and Measuring Economic Resilience. In L. Briguglio, G. Cardigan E. J. Kisanga (eds), *Building the Economic Resilience of Small States.* Malta: Islands and Small States Institute and London: Commonwealth Secretariat, 265-87.

Briguglio, L., Cordina, G., Farrugia, N., & Vella, S. (2008). Economic Vulnerability and Resilience. *UNU-WIDER,* Research Paper*, 55*.

Calem, P. S., Gillen, K., & Wachter, S. M. (2003). *Neighborhood Patterns of Subprime Lending: Evidence from Disparate Cities.* Philadelphia, Pennsylvania: Institute for Law & Economics, University of Pennsylvania.

Calem, P., Gillen, K., & Wachter, S. (2004). The Neighborhood Distribution of Subprime Mortgage Lending. *The Journal of Real Estate Finance and Economics*, *29*.

Chapple, K., & Lester, B. (2007). Emerging Patterns of Regional Resilience. MacArthur Foundation, Institute for Urban and Regional Development, University of California, Berkeley.

Chinitz, B.J. (1961). Contrasts in agglomeration: New York and Pittsburgh. *American Economic Review*. *51*, 279-289.

Christopherson, S., and Clark, J. 2007. Power in Firm Networks: What It Means for Regional Innovation Systems. *Regional Studies*, *41*, 1223-1236.

Clauretie, T.M. (1988). Regional Economic Diversification and Residential Mortgage Default Risk. *Journal of Real Estate Research*, *3*(1), 87-97.

Dale-Johnson, D. (1982). An alternative approach to housing market segmentation using hedonic price data. *Journal of Urban Economics*, *11*(3), 311–332.

Davis, M. A. (2010). Reflections on the Foreclosure Crisis*. Land Lines*, Lincoln Institute of Land Polity. Retrieved on August 10, 2010 from http://www.lincolninst.edu/pubs/PubDetail.aspx?pubid=1821&URL=Reflections-on-the-Foreclosure-Crisis&Page=4

Dissart J. C. (2003). Regional economic diversity and regional economic stability: research results and agenda. *International Regional Science Review*, *26*, 423-446.

Duque, J.C., Ramos, R., Surinach, J. (2007). Supervised Regionalization Methods: A Survey. *International Regional Science Review*, 30,195–220.

Duque, J.C., Anselin, L., Rey, S.J. (2012).The Max-p-Regions Problem. *Journal of Regional Science*, 999. (Forthcoming)

Duval, R., Elmeskov, J., and Vogel. L. (2007). *Structural Policies and Economic Resilience to Shocks*. Economics Department Working Paper 567. Paris: Organization for Economic Cooperation and Development.

Edmiston, K. D. (2009). Characteristics of high foreclosure neighborhoods in the Tenth District. *Economic Review*, (Q II), 51–75.

Eichholtz, P.M. A. (1995). Regional economic stability and mortgage default risk in the Netherlands. *Real Estate Economic*, *23*(4), 421-439.

Ekogen. (2008). Assessing Economic Resilience in Local Economies. Glasgow, Scotland. Retrieved March 27, 2011, from Ekogen Web site: http://www.ekosgen.co.uk/index.html

Essletzbichtler, J. (2005). Diversity, stability and regional growth in the U.S. (1975–2002). Papers in *Evolutionary Economic Geography*, Number 03.13. Utrecht University Urban and regional Research Centre, Utrecht.

Frenken, K., Van Oort, F. G., Verburg, T., & Boschma, R. A. (2005). Variety and regional economic growth in the Netherlands. *Papers in Evolutionary Economic Geography*, *5*, 1–58.

Galster, G. (1996). William Grigsby and the analysis of housing submarkets and filtering. *Urban Studies*, *33*, 1797-1806.

Garcia, R. (2003). Residential Foreclosures in City of Buffalo, 1990 2000. *Federal Reserve Bank of New York*, 1–74.

Garmestani, A.S., Allen, C.R., Mittelstaedt, J.D., Stow, C.A., and Ward, W.A. (2006). Firm size diversity, functional richness and resilience. *Environment and Development Economics*, *11*, 533–551.

Gerardi, K. S., & Willen, P. S. (2008). *Subprime mortgages, foreclosures, and urban neighborhoods* (Discussion Paper 08-06). Boston, MA: Federal Reserve Bank of Boston.

Gilchrist, D. A. and St. Louis. L. V. (1991). Directions for diversification with an application to Saskatchewan. *Journal of Regional Science, 31*, 273-89.

Glaeser, E. L., Scheinkman,J. A., and Shleifer,A. (1995). Economic Growth in a Cross-Section of Cities. *Journal of Monetary Economic*, *36*, 117-43.

Glaeser, E., L., Kallal, H. D., Scheinkman, J. A., and Shleifer, A. (1992). Growth in Cities. *Journal of Political Economy*, *100*, 1126-52.

Goldstein, I. (2008). Subprime lending, Mortgage foreclosure and Race: How Far Have We Come and How Far Have We to Go? .The Kirwan Institute for the Study of Race and Ethnicity, The Ohio State University.

Goodman, A. and T. Thibodeau. (2003). Housing Market Segmentation and Hedonic Prediction Accuracy. *Journal of Housing Economics*, *12*(3), 181–201.

Goodman, A. C. (1978). Hedonic Prices, price indices and housing markets. *Journal of Urban Economics*, *5*, 471-484.

Goodman, A.C. (1981). Housing submarkets within urban areas: definitions and evidence. *Journal of Regional Science*, *21*, 175–185.

Gottlieb, P. D., and Fogarty, M. (2003). Educational Attainment and Metropolitan Growth. *Economic Development Quarterly*, *17*, 325-36.

Grigsby, W. (1963). *Housing Markets and Public Policy*. Philadelphia: University of Pennsylvania Press.

Grover, M., Laura, S., & Richard, T. (2007). *Targeting Foreclosure Intervention: An Analysis of Neighborhood Characteristics associated with High Foreclosure rates in two Minnesota Counties.* Minneapolis: Federal Reserve Bank of Minneapolis, Community Affairs Department.

Hall, B. H. (1987). The Relationship between Firm Size and Firm Growth in the US Manufacturing Sector. *The Journal of Industrial Economics*, *35*(4), 583-600.

Harrison, B., Kelley, M., and Gant, J. (1996). Specialization versus Diversity in Local Economies: The Implications for Innovative Private-Sector Behavior. *Cityscape*, *2*, 61-93.

Hendershott, P., Hendershott, R., & Shilling, J. (2010). The Mortgage Finance Bubble: Causes and Corrections. *Journal of Housing Research*, *19*(1), 1–16.

Henderson, V. (2003). Marshall's Scale Economies. *Journal of Urban Economics*, *53*, 1-28.

Herbert, C. E. (2010). *Report to Congress on the Root Causes of the Foreclosure Crisis*. DIANE Publishing.

Hill, E., St. Clair, T., Wial, H., Wolman, H., Atkins, P., Blumenthal, P., et al. (2010). *Economic Shocks and Regional Economic Resilience*. Washington, DC: The George Washington Institute of Public Policy.

Hill, E. W., Wial, H. and Wolman, H. (2008). *Exploring Regional Economic Resilience*. Working Paper 2008–04, Institute of Urban and Regional Development, University of California, Berkeley.

Holling, C. S. (1973). Resilience and stability of ecological systems. *Annual review of ecology and systematics*, *4*, 1–23.

Hoover, E. M., & Giarratani, F. (1984). *Regional economics*. Alfred A Knopf, New York.

Immergluck, D., & Smith, G. (2005). Measuring the Effects of Subprime Lending on Neighborhood Foreclosures: Evidence from Chicago. *Urban Affairs Review*, 362-89.

Immergluck, D. (2008). From the Subprime to the Exotic: Excessive Mortgage Market Risk and Foreclosure. *Journal of the American Planning Association*, 59-76.

Immergluck, D. (2008). *Community Response to the Foreclosure Crisis: Thoughts on Local Interventions*. Retrieved March 31, 2011, from The Federal Reserve Bank of Atlanta: http://www.frbatlanta.org/filelegacydocs/dp_0108.pdf.

Immergluck, D. (2009). The Foreclosure Crisis, Foreclosed Property, and Federal Policy: Some Implications for Housing and Community Development Planning. *Journal of the American Planning Association*, 406 — 423.

Immergluck, D. (2010). Neighborhoods in the Wake of the Debacle: Intrametropolitan Patterns of Foreclosed Properties. *Urban Affairs Review*, *46*, 3-36.

Jackson, R. W. (1984). An evaluation of alternative measures of regional industrial diversification. *Regional Studies*, *18*, 103-12.

Johnson, M., Turcotte, D., and Sullivan, F. (2010). What foreclosed homes should a municipality purchase to stabilize a vulnerable neighborhood? *Network Spatial Economics*, *10*(3), 363-388

Jones, C., Leishman, C. and Watkins, C. (2004). Intra-urban migration and housing submarkets: theory and evidence, *Housing Studies*, *19*, 269–283.

Jones, C., Leishman, C. and Watkins, C. (2005). Housing market processes, urban housing submarkets and planning policy, *Town Planning Review*, *76*(2), 33–52.

Jones, C., & Watkins, C. (2009). *Housing Market and Planning Policy*. Ames, Iowa: Wiley-Blackwell.

Kaplan, D. H., & Sommers, G. G. (2009). An Analysis of the Relationship between Housing Foreclosure, Lending Practices, and Neighborhood Ecology: Evidence from a Distressed County. *The Professional Geographer*, 101-120.

Kirchhoff, B.A. and Phillips, B.D. (1989). The effect of firm formation and growth on job creation in the United States. *Journal of Business Venturing*, *3*, 261–272.

Kolko, J., and Neumark, D. (2010). Does Local Business Ownership Insulate Cities from Economic Shocks?. *Journal of Urban Economics*, *67*, 103-115.

Kort J. R. (1981). Regional economic instability and industrial diversification in the U.S. *Land Economics, 57*, 596-608.

Lehnert, A., and M. Grover. (2008). New data analysis helps identify future foreclosure trouble spots. In *Community Dividend*. Minneapolis, MN: Federal Reserve Bank of Minneapolis. Retrieved on March 21, 2010 from http://www.minneapolisfed.org/publications_papers/pub_display

Leinberger, C.(2008). The Next Slum? *Atlantic Monthly*, March. Retrieved on March 21, 2010 from http://www.theatlantic.com/doc/200803/subprime

Lipnick, M. D. (1987). *Business applications of shift share analysis*. University of Texas at Austin.

Maclennan, D., & Tu, Y. (1996). Economic perspectives on the structure of local housing systems. *Housing Studies* , 387-406.

Malizia, E. E. and Ke, S. (1993). The influence of economic diversity on unemployment and stability. *Journal of Regional Science*, *33*, 221-235.

Markusen, A. (1985). *Profit Cycles, Oligopoly, and Regional Development*. Cambridge: MIT Press.

Mayer, C., & Pence, K. (2008). *Subprime mortgages: What, Where, and to Whom?* Washington, D.C: Finance and Economics Discussion Series Divisions of Research & Statistics and Monetary Affairs Federal Reserve Board.

Mayer, C., Pence, K., & Sherlund, S. M. (2009). The rise in mortgage defaults. *Journal of Economic Perspectives*, *23*(1), 27–50.

Michaels, R. G., & Smith, V. K. (1996). Market segmentation and valuing amenities with hedonic models: the case of hazardous waste sites. *Estimating economic values for nature: methods for non-market valuation*, *28*, 319.

Mueller, E. (2006). *A Study of Residential Foreclosure in Texas.* Austin, Texas: TDHCA Division of Policy and Public Affairs.

Mueller, G. R. (1993). Refining Economic diversification strategies for real estate portfolios. *Journal of Real Estate Research*, *8*, 55-68.

Munro, M. (1986). Testing for segmentation in the private housing market in Glasgow. *Centre for Housing Research,* Discussion Paper No, 8.

Newman, K., & Wyly, E. K. (2004). Geographies of mortgage market segmentation: The case of Essex County, New Jersey. *Housing Studies*, *19*(1), 53–83.

Ong, P., & Pfeiffer, D. (2008). *Spatial Variation in Foreclosures in Los Angeles.* Los Angeles, California: Ziman Center, University of California.

Pedersen, C., & Delgadillo, L. (2007). Residential Mortgage Default in Low- and High-Minority Census Tracts. *Family and Consumer Sciences Research Journal*, 374-391.

Pendall, R., Foster, A. K., and Cowell, M. (2007). *Resilience and Regions: Building Understanding of the Metaphor*. Prepared for Building Resilient Regions, Research Network Sponsored by the MacArthur Foundation, Institute for Urban and Regional Development, University of California, Berkeley.

Perkins, K. L. (2009). *The Geography of Foreclosure in Contra Costa County, California.* San Francisco: Federal Reserve Bank of San Francisco.

Pryce, G. and Evans, G. (2007) Identifying Submarkets at the Sub-Regional Level in England, London: Department of Communities and Local Government.

Quigley, J. M. (1998). Urban Diversity and economic growth. *Journal of Economic Perspectives*, *12*(2), 127-138.

Rey, S.J., L. Anselin, D.C. Folch, M.L. Sastre-Guitierrez and L. Interlante (2011). Measuring spatial dynamics in metropolitan areas. *Economic Development Quarterly*, *25*, 54–64.

Rose, A. (2007). Economic resilience to natural and man-made disasters: Multidisciplinary origins and contextual dimensions. *Environmental Hazards*, *7*(4), 383–398.

Rose, A., & Liao, S. Y. (2005). Modeling Regional Economic Resilience to Disasters: A Computable General Equilibrium Analysis of Water Service Disruptions. *Journal of Regional Science*, *45*(1), 75–112.

Rothenberg, J., Galster G. C., Butler, R. V. and Pitkin, J. R. (1991). *The Maze of Urban Housing Markets*. Chicago: University of Chicago Press.

Shaffer, S. (2002). Firm size and economic growth. *Economics Letters*, *76* (2), 195-203.

Sherwood-Call, C. 1990. Assessing regional economic stability: A portfolio approach. *Economic Review* (Federal Reserve Bank of San Francisco), 17-26.

Smith, S. M., and Gibson, C.M. (1988). Industrial diversification in nonmetropolitan counties and its effect on economic stability*. Western Journal of Agricultural Economics, 13*, 193-201.

Straszheim, M. R. (1975). *An econometric analysis of the urban housing market.* NBER Books*.*

Tobler, W. R. (1970). A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography, 46,* 234-240.

Tu, Y. (1997). The Local Housing Submarket Structure and its Properties. *Urban Studies*. 34, 337-353.

You, J.I., (1995). Small firms in economic theory. *Cambridge Journal of Economics*, *19* (3), 441–462.

Wagner, J. E. and Deller. S.C. (1998). Measuring the Effects of Economic Diversity on Growth and Stability. *Land Economics*, *74*, 541-556.

Watkins, C. (2001). The definition and identification of housing submarkets*. Environment and Planning A*, *33*, 2235-2253.

Williams, J. E. (1996). Real Estate Portfolio Diversification and Performance of the twenty largest SMSAs. *Journal of Real Estate Portfolio Management*, 2(*1*), 19-30.